This is a reproduction of a library book that was digitized by Google as part of an ongoing effort to preserve the information in books and make it universally accessible.

Google™ books

https://books.google.com

AF MANUAL 101-8

FUNDAMENTALS OF
ELECTRONICS

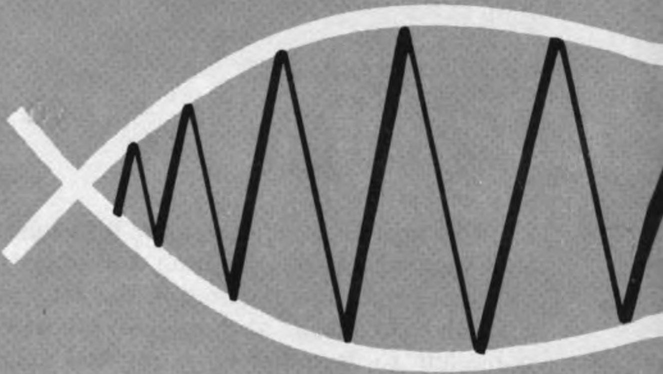
DEPARTMENT OF THE AIR FORCE

BEHREND COLLEGE
LIBRARY
Erie, Pennsylvania

THE PENNSYLVANIA STATE UNIVERSITY · 1855 ·

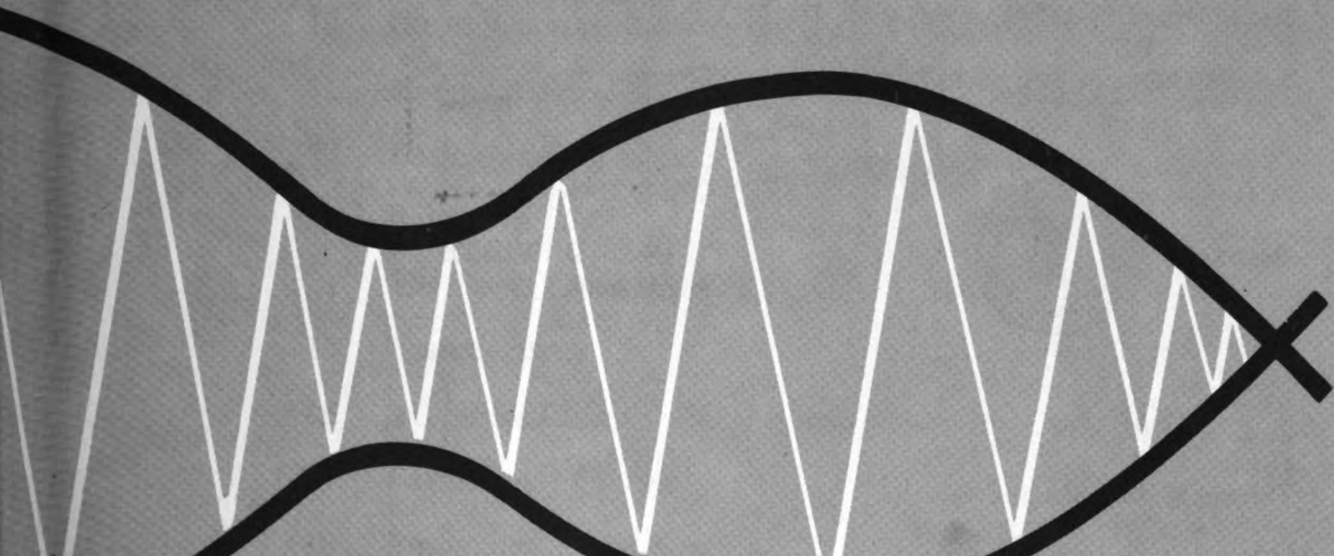
THE PENNSYLVANIA STATE
UNIVERSITY LIBRARIES

AF MANUAL 101-8



FUNDAMENTALS OF

ELECTRONICS



THE PENNSYLVANIA STATE UNIVERSITY
COMMONWEALTH CAMPUSES LIBRARIES
BEHREND

DEPARTMENT OF THE AIR FORCE

UG633
.A5763
no. 101-8

AIR FORCE MANUAL
NUMBER 101-8*

DEPARTMENT OF THE AIR FORCE
WASHINGTON, D. C.
1 JULY 1957

FOREWORD

This manual covers the basic principles of electronics the aircraft navigator needs for an understanding of the operation and adjustment of radar equipment. It develops electronic principles from those required in the analysis of simple DC circuits to those involved in complex radar circuits. It also provides the navigator with an electronics reference text that will enable him to use radar operating and maintenance instructions more intelligently.

This manual covers material which includes DC circuits and principles, magnetism, AC circuits and principles, electron emission and electron tubes, power supplies, amplifiers, oscillators, and the transmission and reception of radio waves. Radar circuits such as transients, square-wave generators, blocking of oscillators, sweep generators, clampers, and clippers are then covered. After a discussion of radar principles, the following topics are presented: microwave oscillators, transmission lines, wave-guides, and antennas. A transmitter channel, a receiver channel, and an indicator channel of a typical radar set are then described. A chapter on synchros and one on mathematics are also contained in the manual. The Appendix includes tables of common logarithms, tables of trigonometric functions, and a section on electronic schematic symbols.

Recommendations and suggestions for the improvement of this manual are encouraged and should be forwarded to the Director of Personnel Procurement and Training, DCS/P, Headquarters USAF, Washington 25, D. C.

BY ORDER OF THE SECRETARY OF THE AIR FORCE:

THOMAS D. WHITE
Chief of Staff

OFFICIAL:

J. L. TARR
Colonel, USAF
Director of Administrative Services

*This manual supersedes ATC Manual 101-2, Oct 55

Copyright material contained in this publication has been reproduced by special permission.

OPI: AFPTR
DISTRIBUTION: F

This training manual is published by Air Training Command in accordance with AFR 5-5

REPRINTED WITH CHANGE A, 25 JAN 1963, INCLUDED

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington 25, D. C. — Price $4.00

contents

chapter	title	page no.
1	Why Electronics for the Navigator	1
2	Static Electricity	3
3	Electricity in Motion	9
4	Introduction to Circuits	17
5	Series Circuits	27
6	Parallel Circuits	33
7	Series-parallel Circuits	37
8	Magnetism	43
9	Electromagnetic Induction and Generators	51
10	Direct Current Motors	61
11	Meters	69
12	Basic Concepts of Alternating Current	81
13	Inductors	87
14	The Basic Capacitor	99
15	Transformers	107
16	AC Series Circuits	115
17	Parallel AC Circuits	127
18	Resonance	135
19	Electron Emission and the Diode	149
20	Triodes	159
21	Tetrodes and Pentodes	173

chapter	title	page no.
22	Special Purpose Tubes	179
23	Cathode-ray Tubes	189
24	Introduction to Radio Transmission and Reception	197
25	Power Supplies	203
26	Principles of Amplification	215
27	Principles of Oscillation	233
28	Principles of Transmission	251
29	Principles of Radio Reception	267
30	Transients and Time Constants	283
31	Blocking Oscillators	301
32	Nonsinusoidal Generators and Waveshaping Circuits	309
33	Basic Radar Principles	323
34	Microwave Oscillators	337
35	Transmission Lines	353
36	Waveguides, Cavity Resonators and Duplexers	381
37	Radiation and Antennas	411
38	A Radar System	433
39	Synchros and Servomechanisms	471
40	Technical Orders	505
41	Mathematics for Navigators	507
	Appendix	537
	Circuit symbols	553
	Index	557

Chapter 1

Why Electronics for the Navigator

The Air Force wants you to become a trained specialist in the challenging field of electronics. As a navigator you will operate and perform in-flight maintenance on complex electrical and electronic equipment which has greatly aided the Air Force to become a big, modern organization.

None of the equipment can keep itself cared for and repaired. The proper functioning of each piece of equipment depends completely upon the know-how of the operator and repairman assigned to it. Consequently, without the services of the operator and repairman, all Air Force electronic equipment would be just so much scrap metal.

While on the job as an Air Force navigator, you will find yourself in a world of men who are doing a highly skilled task — and they are proud of it! They are proud of the fact that their skill is essential to destroying the enemy and making it possible to return home safely in any weather.

Since success in present-day flight operations is largely dependent upon the effective use of radio and radar equipment, we may say that, next to the atomic bomb, the perfection of radar during World War II was the most important scientific development in Air Force planning.

Although the development of radar was one of the best kept secrets of the Second World War, it is not as new nor radical as is commonly imagined. Radar was a natural outgrowth of intensive radio research over a period of many years. During the years before the war, two types of radio equipment from which radar took many of its basic ideas were developed. The first was the compact type of direction-finding mechanism which used directional antennas. The second was television which used the cathode-ray tube as an indicator.

One of the first observations of radio echoes was made in the United States in 1922 by Dr. A. Hoyt Taylor of the Naval Research Laboratory. Dr. Taylor observed that a ship passing between a radio transmitter and a radio receiver reflected some of the waves back toward the transmitter.

Further research led in 1934 to the determination of range by a single set — the chief accomplishment of radar. A pulse radar set, for aircraft detection, was demonstrated at the Naval Research Laboratory at Anacostia in June 1936, and afloat on the destroyer Leary in April 1937.

The development and use of airborne equipment received little encouragement until, in April 1940, radar enthusiasts in the RAF Coastal Command sighted the mission German prison ship *Altmark* in a Norwegian Fjord with early design ASV (Air to Surface Vessel) equipment. The Altmark incident disclosed the usefulness of airborne radar, and rapid research in this type of equipment soon led to the perfection of night-fighter radar. The requirements of airborne equipment were met by higher frequencies and led to the development of microwave radar, perhaps the greatest advancement in radar made during the war. The higher precision of microwave radar allowed it to be employed for such offensive measures as bombing and gunlaying which were greatly in demand after the Allied Air Forces had assumed an offensive role.

Research on airborne radar led to a remarkable series of developments which resulted in the adoption of devices that identified friendly aircraft by automatic coded replies accompanying the reflected echoes. These devices are known as IFF (Identification, Friend, or Foe).

Further developments of radar led to the construction of associated beacons as navigational aids. The navigational possibilities of radar were quickly recognized, and amazing progress in the field of airborne, navigational equipment resulted. Future trends in electronic warfare indicate that rapid progress is being made in improved navigational aids, identification and gunlaying equipment, guided missiles, search equipment, and electronic countermeasures.

It should be apparent that there is a definite need for you to understand electronic fundamentals in order that you may better function in your special field. Modern radio and radar equipment is based on a common core of electricity and electronics; consequently, a man who has a firm grasp of electronic fundamentals is a definite advantage to his teammates. You may competently operate highly complex electronic equipment, often in times of crisis, provided that you have been thoroughly grounded in electronic fundamentals. Advanced radio and radar instruments are simply outgrowths of elementary electronic circuits.

As a navigator your work covers the operation and in-flight maintenance of radar equipment and all types of electronic devices used by the Air Force. A partial listing of some navigator jobs would include such specialities as Shoran and Loran, all weather interception, the K-system, navigation, bombing, and radar.

In addition to being assigned responsible and attractive duties in the Air Force, the individual navigator can look forward to a further advantage. Present plans call for commissioned navigators in excess of the number required to fill crew positions, thus affording navigators the opportunity to branch out into other occupational fields. The type of training offered in the Navigator Training Program provides an excellent background for advancement in such varied fields as Intelligence, Communication-Electronics, Armament, Cartography, Production Procurement, and Education.

The navigator's training may be of value in civilian as well as in military life. Consider the dollars and cents value of your Air Force training as compared to similar civilian training. Few, if any, civilian schools engaged in electricity and electronics training have the investment in equipment, training aids, and laboratory facilities that your Air Force schools have. Your effort in the school, whether applied to study or practice or to the development of your character, also increases your value as a citizen and a worker in civilian life. The more you learn as a navigator and the better job you do in the Air Force, the higher you will advance in the Air Force, and the more success you will have, both in the Air Force and out.

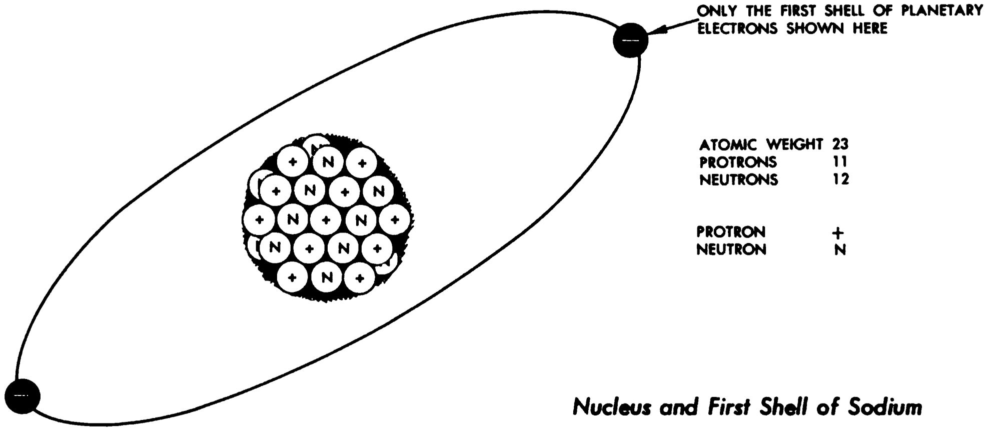
Chapter 2

Static Electricity

The chair you are sitting on, this paper in your hands, the very words you are reading, YOU — are MATTER. Radar sets, their component units, the wires, and tubes, in fact everything that goes to make up a set, are examples of matter. Technically, matter is anything that *occupies space* and *has weight* and *mass*. Anything so all-inclusive is of great interest to anyone; however, to you, as a navigator, the structure of matter is of primary importance. From the study of matter, static and dynamic electricity assume significance and become understandable. Matter is found to exist in three states: solid (iron, wood, etc.), liquid (water, alcohol, etc.), and gas (air, oxygen, etc.).

STRUCTURE OF MATTER

Curiosity!! The child rips his toy apart, the sixteen-year-old boy disassembles and rebuilds his jalopy, the scientist attempts to answer the question, "Why?" This healthy curiosity about matter occupies man's mind now as it has ever since time began. The Ancient Greeks were convinced that all physical things consisted of four *elements:* air, earth, fire, and water. To them materials were different because they were composed of varying relative amounts of these four elements. As the centuries passed and scientific methods were developed, the four Greek elements were broken down into more and more "indivisible" substances which still retain the name *element*. At least 98 such elements have been isolated. An *element* (such as iron, gold, or oxygen) is a substance which cannot be reduced to a simpler substance by chemical means. All matter consists of elements or combinations of elements. These combinations are of two types: compounds and mixtures. The *compound* (such as water or salt) consists of two or more elements which are *chemically* united. The *mixture* (such as air or milk) is a *physical* binding of several elements or compounds.



Nucleus and First Shell of Sodium

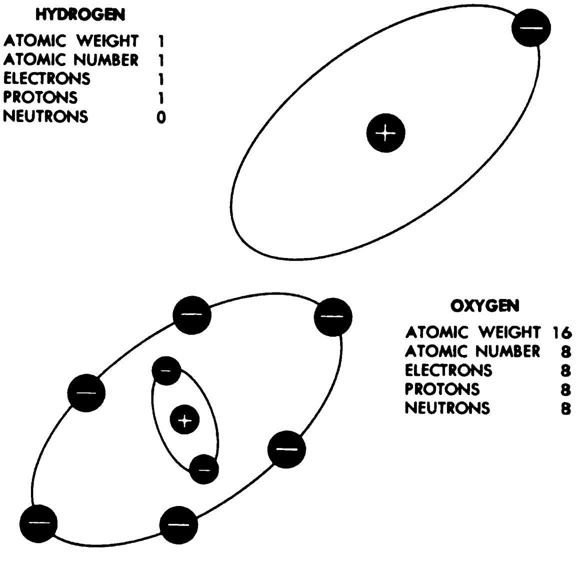
STRUCTURE OF THE ATOM

Compounds and elements can be reduced to smaller and smaller parts until only a minute particle of matter remains. The smallest particle of matter that still retains all the characteristics of the original element or compound is the *molecule*. The molecule, in turn, consists of atoms, chemically united. The smallest particle of the element is the *atom*.

The atom itself has been found to consist of many smaller particles — three of which are the *proton*, *electron*, and *neutron*. The diameter of an atom is estimated in millionths of an inch while the electron is even smaller (trillionths of an inch). It is not difficult to see, therefore, that the dot at the end of this sentence is composed of millions of atoms. If the dot were magnified to the size of the earth, such atom would be about the size of a baseball and the electrons would still be invisible.

As shown in the illustration above, the *nucleus* is a comparatively compact group of protons and neutrons near the center of the atom. The electrons of the atom revolve around its nucleus in orbits spaced at varying distances from the nucleus. Because of their resemblance to the planets of the solar system, these electrons are often referred to as *planetary* or *orbital* electrons. Atoms of one element differ from atoms of other elements in the number and distribution of protons, electrons, and neutrons. This is shown in the illustration, Atoms of Hydrogen and Oxygen.

In order to illustrate the breakdown of matter, let us take sea water, obviously a form of matter. It is a *mixture* consisting chiefly of salt and water, which can be separated from each other by physical means (distillation). The water is a *compound*, consisting of one part of oxygen to two parts hydrogen, which can be separated only by means of a chemical reaction. One *molecule* of water consists of one *atom* of oxygen, chemically combined with two atoms of hydrogen. Now consider the atom of oxygen. It has



Atoms of Hydrogen and Oxygen

eight *protons* and eight *neutrons* in its *nucleus,* and eight planetary electrons revolve around the nucleus. Note that the number of planetary electrons is equal to the number of protons — this is true of all neutral atoms.

CHARGED BODIES

Thales, a Greek philosopher (about 600 B. C.), discovered that when amber was rubbed, it attracted certain light materials. You can duplicate Thales' findings by using rubber in place of amber. Obtain two or three bits of paper (about ¼ inch square) and a rubber comb or other hard rubber object. Rub the comb briskly with wool (the sleeve of your coat may do) and try picking up the pieces of paper. You have produced electricity!

In the seventeenth century further experiments were performed to determine the cause of this phenomenon of attraction. These experiments resulted in the discovery that many materials (such as glass, rubber, and sealing wax) show characteristics similar to amber. One experiment shows that when a rod of glass is rubbed with silk and suspended to swing freely, it will swing away from another rubbed glass rod approaching it (repulsion). However, if a suspended glass rod is approached by a rubber rod that has been rubbed with a cat's fur or wool, the glass rod swings toward the rubber (attraction). Evidently, the condition of the rubber and the condition of the glass were changed by being rubbed, for unrubbed materials show neither attraction nor repulsion. The rubbed rubber and glass are said to be *charged.* Arbitrarily, the charge on the rubber is called negative (−) while that on the glass is called positive (+). From this experiment we arrive at the first and one of the most fundamental laws of electricity: *like charges repel each other and unlike charges attract.* Since these charges are not in motion, they are *static* (at rest). The study of static electricity is the study of electrical charges at rest.

Why do the rubber and glass become charged when rubbed? Why are they charged differently? It is apparent that rubbing the glass with silk causes some change to occur. Experiments indicate that some of the planetary electrons are removed from the glass by the silk, leaving the glass with a deficiency of electrons. If the glass is to be positive (the assumption of the previous paragraph) then the *electron* is *negative,* because only by the removal of negative bodies can a neutral body be made positive. *Each electron is a unit negative charge.* Other experiments show that the proton has greater weight then the electron (1840 times as much) and is of equal but opposite charge. *Each proton is a unit positive charge.* The neutron combines one electron and one proton and therefore has no charge (is neutral).

IONS AND IONIZATION

Not all atoms are neutral!! When the number of electrons is not equal to the number of protons, the atom is an *ion.* When an atom has an excess of electrons, because the number of electrons exceeds the number of protons, the atom is said to be a negative ion. An atom with a deficiency of electrons is a *positive* ion. The process of subjecting the neutral atom to the removal or addition of electrons is known as *ionization.*

The following facts have been shown:

Neutral atoms are composed of equal numbers of electrons and protons.

With the proper force, it is possible to add electrons to the *outer* orbit of an atom.

Conversely, it is possible to take electrons out of the *outer* orbit of an atom. These electrons are referred to as *free electrons.*

The forces required to add electrons to the outer orbit of an atom or to take them out may be supplied by charged bodies.

A negatively charged body is one having an excess of electrons, while a positively charged body is one having a deficiency of electrons.

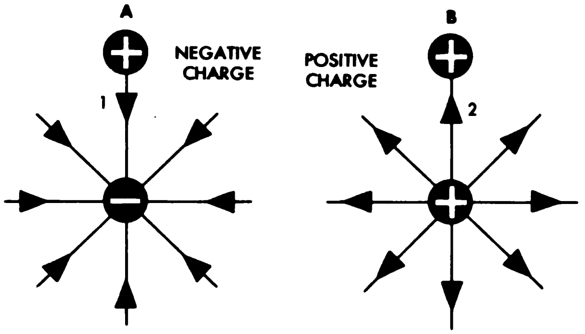
CHARGES EXERT FORCE

If the nucleus is positive and if the planetary electrons are negative, what prevents the force of attraction from causing a crash? Electrons will not crash into the nucleus for the same reason that the moon does not crash into the earth or the earth into the sun. Picture a boy twirling a weight tied to a string. As long as the string holds out and as long as the boy continues the twirling, the weight will travel in an orbit. The weight has two forces acting on it: (1) *the force of motion tending to make it fly off the string,* and (2) *the pull of the string on the weight toward the boy.* The resultant force keeps the weight moving in an orbit. In the atom the situation is similar. The electron (weight) attempts to fly off into space because of its motion, but the electrostatic attraction (string) exerted by the nucleus (boy) keeps it in its orbit.

The repulsion between the glass rods and the attraction between the glass and rubber prove that a force exists between charged bodies. How much force? What affects this force? In the 18th century, the French physicist Coulomb measured the force (attraction or repulsion) between two charged bodies and found that the force increased as the charges on the bodies increased and decreased as the distance between them increased. Coulomb's law which expresses this relationship is expressed as follows: The force between two charged bodies is directly proportional to the product of the charges and inversely proportional to the square of the distance between them.

FIELDS OF FORCE

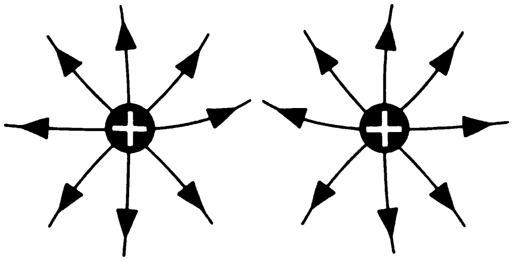
The radar oscilloscope, the television picture tube, the radio tube, and countless other electronic devices employ static electricity to affect moving charged bodies. What is there about a charged body that influences a moving charge? Since Coulomb's law is always true, any charged body in the vicinity of another charged body will have a force exerted on it. We express this fact by saying that *fields of force* exist about all charged bodies. Because fields of force are invisible, a method had to be devised to represent them on paper. As shown in the illustration Fields of Force around Isolated Charges, the field of force about a charged body is represented by drawing lines with arrows to show the path and direction a *unit positive charge* would follow. These lines are called lines of force and represent the intensity and direction of electrostatic fields. The intensity is indicated by the spacing of the lines, the direction by arrows. An examination of the illustration will show that a unit positive charge placed at point A will take the path shown by arrow 1 toward the negative body, while the same charge placed at point B would move away from the positive body as indicated by arrow 2.



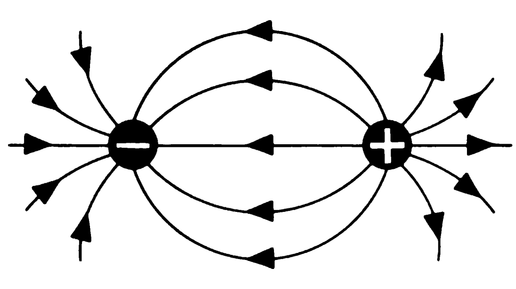
Fields of Force around Isolated Charges

In most practical applications, the electrostatic field is created by the presence of two or more charged bodies or surfaces. The effect of complex fields is reduced to a comparatively simple situation when lines of force are used. As in the case of single charged bodies, a line of force represents the path a unit positive charge takes if placed in the field. You must remember that lines of force do not exist but are representations of forces that *do* exist. The actual force completely fills the area indicated by lines of force. Examine the illustration below and note that the lines of force bulge and repel each other. A good rule to remember is that *lines of force acting in the same direction tend to repel each other.*

LIKE CHARGES



UNLIKE CHARGES



Fields of Force between Charges

THE ELECTROSCOPE

An electric charge may be detected by its ability to attract light objects, such as pith balls and bits of paper. Any device used for detecting electric charges is called an electroscope.

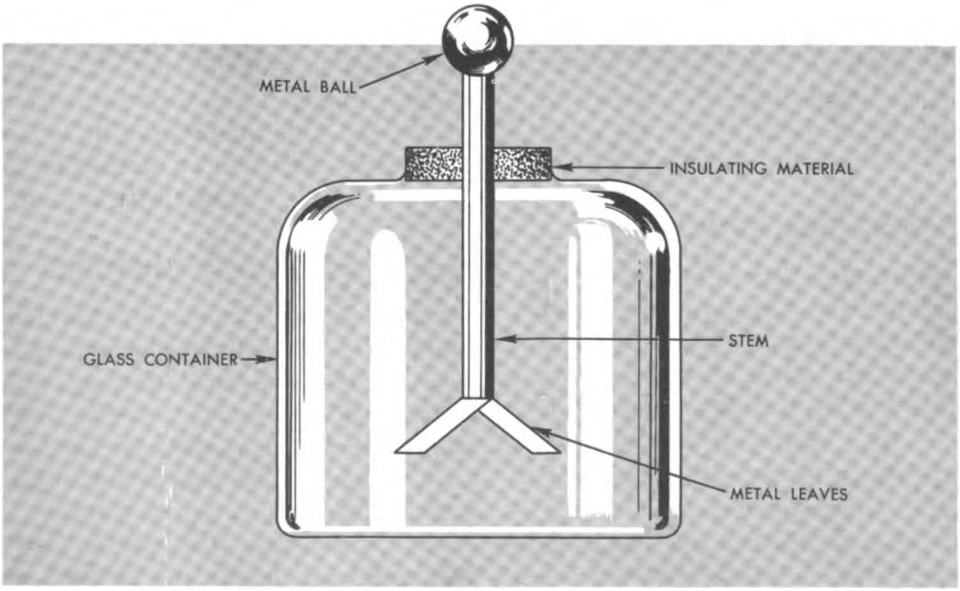
In its simplest form, an electroscope consists of a pith ball hanging on the end of a silk thread. By touching the ball with a body of a known charge, you have an instrument capable of detecting charged bodies and of indicating the type of charge or polarity. If you touch the pith ball with a glass rod which has been rubbed with silk, you charge it positively. Any other charged body which is brought near the pith ball will repel it if the body is positive, or attract it if the body is negative. The force of repulsion or attraction indicates the strength of the field surrounding the charged bodies.

A better and more sensitive device is the leaf electroscope shown in the illustration below. It consists of two thin sheets of metal foil (usually gold or aluminum), called leaves, supported by a wire or stem which passes through a block of sealing wax or insulating material and ends in a metal ball or cap. The leaves are usually sealed in a glass container to prevent air currents and moisture from affecting the instrument. The sensitivity of the instrument depends on how thin and of what material the leaves are made.

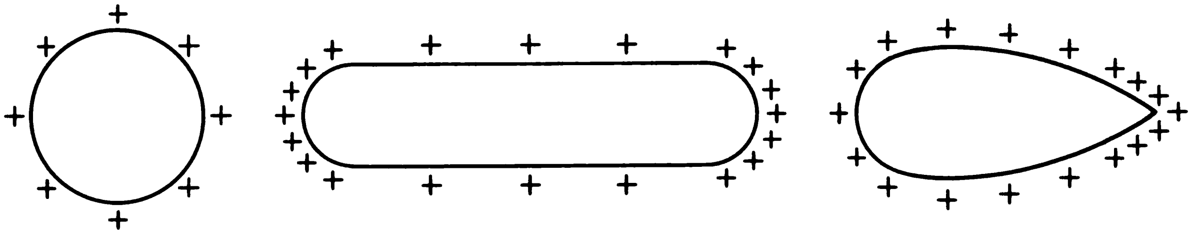
If the ball receives either a positive or a negative charge, it will react on the leaves and spread them apart. The spreading of the leaves is due to like charges being on each leaf. After a charge of positive electricity has been placed on the leaves, the spread of the leaves increases if a positively charged body is brought near the ball. On the other hand, a negatively charged body brought near the ball will cause a decrease in the spread.

DISTRIBUTION OF CHARGES ON OBJECTS

The surface density of charge of an object depends upon the shape of the object. This can be shown experimentally in the following way. By touching the charged object with a proof plane (a tiny metal plate with an insulated handle), part of the charge of the object is removed. If the proof plane is brought in contact with an electroscope, the amount of rise of the gold leaves of the electroscope can be taken as a measure of the charge on the plane or the charge density of the object. In this way we find that the density of the charge on the outside of a sphere is uniformly distributed. On charged objects, other than spheres, the greatest



Leaf Electroscope



Distribution of Charges

density of charge is found on the part which has the greatest curvature. Thus, if a teardrop-shaped object is charged, the intensity of the electric field will be greatest in the region of the sharp point as shown in the illustration above. Use of this fact is made in the design of lightning rods and of certain spark gaps. The intensity of the electric field at which the spark jumps across the gap is determined by the shape of the electrodes. The sharper the points, the lower the intensity needed for a given separation.

If we take a hollow spherical conductor with a hole in it and explore the inner surface of the object, we find that regardless of the amount of charge on the outer surface, *no charge exists on the inner surface.* This can be shown by inserting a proof plane into the charged sphere and making contact with the inner surface. After removing the proof plane, test it by means of an electroscope. If any charge is present on the inner surface, a part of it will be transferred to the proof plane and the electroscope will indicate the presence of the charge. This experiment always shows no measurable charge.

This property of a closed conductor is the basis of electrostatic shielding; that is, circuit elements are enclosed in metal cans to isolate them from outside electric fields.

SUMMARY

Matter is anything that occupies space and has weight and mass. It consists of elements or combinations of elements. The smallest particle of the element is the atom which has a nucleus composed of protons and neutrons and planetary electrons. An electron is a unit negative charge, and a proton is a unit positive charge.

Static electricity is the study of electrical charges at rest. Like charges repel each other and unlike charges attract each other with a force inversely proportional to the square of the distance between them (Coulomb's law). Fields of force exist around all charged bodies.

Chapter 3

Electricity in Motion

In the last chapter you saw that an excess of free electrons on a body causes it to be negatively charged, while a deficiency of free electrons causes it to be positively charged. If two differently charged bodies are connected by a wire, electrons flow from the body having the greater negative charge to the one having the lesser negative charge. This flow of electrons is an *electric current* and is identical with that in the filament of an electric light bulb, the heating element of an electric water heater, or the field windings of an electric motor.

ELECTRIC CURRENT

The magnitude of current depends upon the number of electrons which flow in a given time. Electrons are so small and the charge on each is so minute that only great numbers of electrons are of any practical consequence in electricity. Approximately 6.3 million million million electrons are referred to as a coulomb, a more practical unit of charge. If *one* coulomb passes by a given point in *one* second, the magnitude of the current flow is one ampere. Current is the rate of flow of electrons and is represented by the capital letter I. The ampere, the unit of current flow, is designated by the abbreviation amp. For reasons of expediency in standardizing measurements, the ampere has been defined as the constant current which will deposit silver in electroplating at the rate of 0.001118 gram per second from a silver nitrate solution under prescribed conditions. This quantity of current is more easily visualized by referring to a 100-watt light, which ordinarily has a current of about one ampere flowing through it.

RESISTANCE AND CONDUCTANCE

A wire used to connect two differently charged bodies serves to conduct current. This ability to conduct current varies according to the substance through which the current flows and depends upon the number of free electrons which the substance possesses.

There is no perfect conductor. Every substance has at least some opposition to the flow of current, and this property is called *resistivity*. The opposition given by a particular object to the flow of current is called the *resistance*, and its symbol is the letter R. Resistance is measured in a unit called an *ohm*, which is represented symbolically by the Greek letter omega (Ω). By the international standard, an ohm is specified as the resistance offered to the flow of current by a column of mercury having a uniform cross section, a length of 106.3 centimeters, and a mass of 14.4521 grams at a temperature of 0° C. The resistance of an electric toaster or an electric iron is about 25 ohms.

A substance in which each atom has only relatively few free electrons is termed a *resistor* and opposes the flow of current. A radio affords many examples of different sizes and types of resistors which will be discussed later.

A substance possessing a relatively great number of free electrons conducts current very readily and is called a *conductor*. The chief purpose of a conductor is to conduct current from one place to another. For example, wires are used to carry current from a generator to a radar set. The ability of a substance to conduct electrons is called its conductivity. The greater the number of free electrons possessed by a material the greater the conductivity.

Some common metals arranged in decreasing order of conductivity (silver being the best conductor) are silver, copper, gold, aluminum, tungsten, zinc, iron, platinum, steel, and tin. Numerically, conductivity is the reciprocal of resistivity. The ability of a particular object to conduct current is called its *conductance*, the unit of which is the *mho* and the symbol of which is G. For example, a piece of copper wire that is about one-tenth of an inch in diameter and is 1000 feet long has a conductance of one mho.

A substance which has so few free electrons that, for all practical purposes, it allows no current to flow is an *insulator*. There is no perfect insulator. However, there are many insulators which allow only a very small current to flow. The rubber around the wire which brings the current to your household electrical appliances is an example of an insulator.

Silver and copper may be classified as conductors; iron, tin, and carbon as resistors, and glass and porcelain as insulators. The greater the number of free electrons, the greater the conductivity and the less the resistance. On the other hand, the fewer the number of free electrons possessed by a substance, the less the conductivity and the greater the resistance.

The Effect of Dimensions on Resistance

Wires which are used to conduct current from one place to another are usually made of copper or aluminum. They are usually insulated by either a rubber compound or a varnished cambric. Wires vary in the value of their resistance according to their lengths and according to their cross-sectional areas. The resistance is directly proportional to the length and inversely proportional to the cross-sectional area. If two similar wires have equal cross-sectional areas, but one is twice as long as the other, the longer wire has twice as much resistance as the shorter one. On the other hand, if two similar wires have equal lengths, but one has one-third as great a cross-sectional area as the other, the smaller wire has three times as much resistance as the larger one.

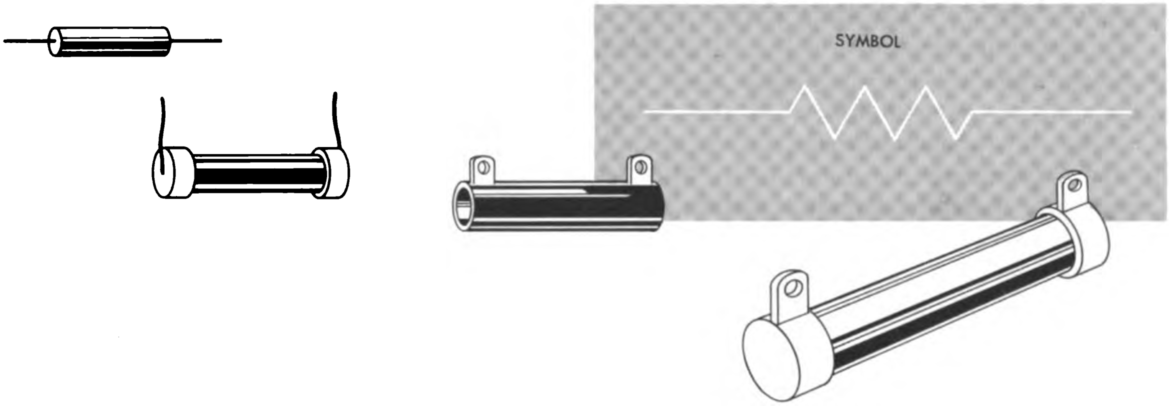
The Effect of Temperature on Resistance

The resistance of some materials, such as metals, increases with temperature. Within practical limits, the change in the resistance of wire is usually proportional to the change in temperature. For example, a given copper wire has a resistance of 100 ohms at 0° C, a resistance of 100.85 ohms at 2° C, and a resistance of 101.7 ohms at 4° C. A substance having this property of an increase in resistance with an increase in temperature is said to have a *positive temperature coefficient.* Carbon is one of the few substances that has a *negative temperature coefficient;* that is, its resistance decreases as the temperature increases.

FIXED AND VARIABLE RESISTORS

Although temperature affects the value of a resistor, a relatively constant value of resistance may be introduced in a circuit by the use of a fixed resistor. Such resistors are used in a circuit where a constant resistance is desirable. The illustration at the top of the next page shows the picture of several types of fixed resistors, together with the corresponding symbol that is used in scientific literature to represent them.

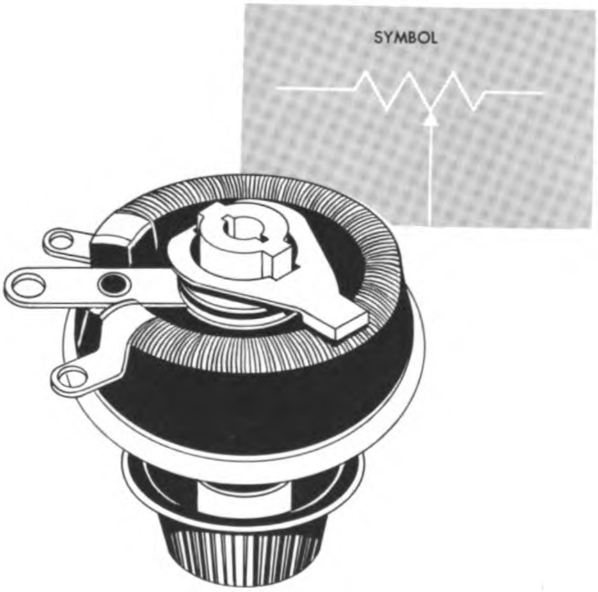
Adjustable resistors are used where it is necessary to change, or adjust, the value of the resistance in a circuit from time to time. In its usual form, the adjustable resistor is wire-wound, and has one or more sliding collars which may be moved along the resistance element to select



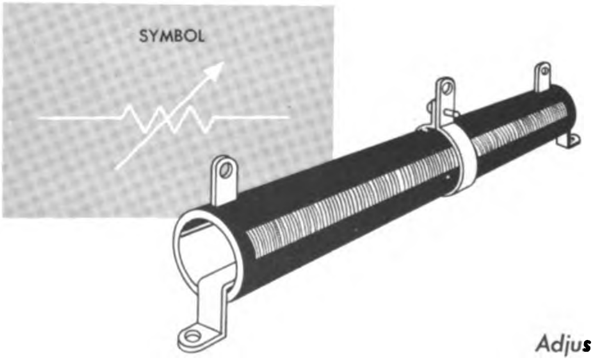
Fixed Resistors and Symbol

the desired resistance value. After the desired value has been selected, the collars are then clamped in place. An adjustable resistor is shown in the illustration below.

Variable resistors are used in a circuit when resistance must be changed frequently. Variable resistors are either of carbon or are of wire-wound construction, depending on the current that is to flow through them. The actual resistance element of a variable resistance is usually circular in shape, and the sliding tap, or arm, which makes contact with it is provided with a knob and a shaft by means of which the resistance can be varied smoothly. Variable resis tors are generally called *rheostates* or *potentiometers,* depending on how they are used. A type of variable resistor is shown in the illustration to the right.



Variable Resistor with Symbol



Adjustable Resistor and Symbol

RESISTOR COLOR CODE

To determine the value of the adjustable or variable resistors, you need only look for a stamp on an exposed surface of the resistor. However, it was found to be impractical to stamp or mark fixed resistors numerically. Manufacturers have therefore devised a system of color coding resistors of the carbon variety.

In the color code system, each number has the following color significance.

Number	Color
0	Black
1	Brown
2	Red
3	Orange
4	Yellow
5	Green
6	Blue
7	Violet
8	Gray
9	White

These colors are used in combinations to represent the value of the resistance that is used. Since it is difficult to make a resistor with exactly a given value of resistance, a color is also put on the resistor to designate the percentage variation a resistor can be expected to have. This percentage variation is known as *tolerance*. The markings for tolerances are as follows:

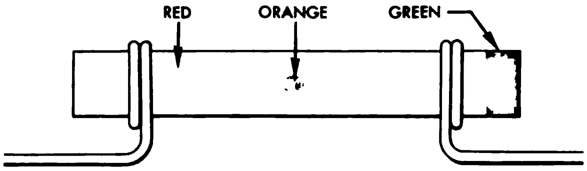
Color	Tolerance
Brown	1%
Red	2%
Orange	3%
Yellow	4%
Gold	5%
Silver	10%
No Marking	20%

The last three tolerances are the most common.

There are two methods of color coding a resistor. The first makes use of the following sequence: body color, end color, and the color of a dot on the resistor. As a memory aid, the color sequence can be associated with the word BED, obtained from the initial letters of each word — body, end, dot. The body color indicates the first digit to be set down, the end (or tip) color indicates the second digit, and the color of the dot represents the number of zeros that must be added to the first two digits. Thus, the resistance in the next illustration is read as follows:

Red	Green	Orange	No color
2	5	000	20% Tolerance

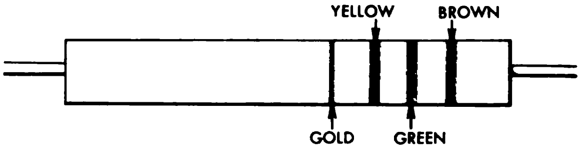
The value of the resistance, then, is 25,000 ± 5,000 ohms, which means that this resistor may have a resistance between 20,000 ohms and 30,000 ohms.



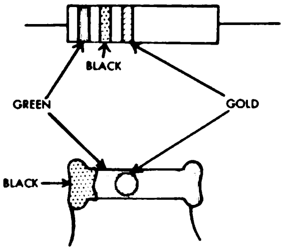
A second method which has been adopted by the Radio Manufacturers Association uses three bands. The colors are read from the nearest tip (end) toward the center. The first band indicates the first digit, the second band the second digit, the third band the number of zeros, and the fourth band the tolerance. In the illustration below, the value of the resistor is read as follows:

Brown	Green	Yellow	Gold
1	5	0000	5% Tolerance

The value of the resistor is 150,000 ± 7,500 ohms.



Occasionally, the need arises for color coding resistors with values below 10 ohms. This is accomplished by making the third band (the third significant figure) a divisor; gold signifies $\frac{1}{10}$ and silver, $\frac{1}{100}$. The figure below illustrates a 5 ohm resistor ±20%.



5 Ohm Resistor ± 20%

POTENTIAL DIFFERENCE

Let us assume that a given charged body has 90 free electrons and another charged body of equal size has 10 free electrons. If these two bodies are connected by a wire, electrons flow from the more negatively charged body to the one having the lesser negative charge. This continues until there are 50 electrons on one body and 50 on the other. At this time, the two bodies are equally charged and no electrons flow along the connecting wire. The electrons move because of the difference in charge between the two bodies. This difference of charge is called a *difference of potential* or *potential difference.*

The unit of potential difference is the volt. A volt is defined as the potential difference which causes one ampere to flow through a resistance of one ohm. The potential difference of an ordinary dry cell, such as a cell of a flashlight battery, is one and one-half volts. The potential difference for most domestic electric service is approximately 115 volts.

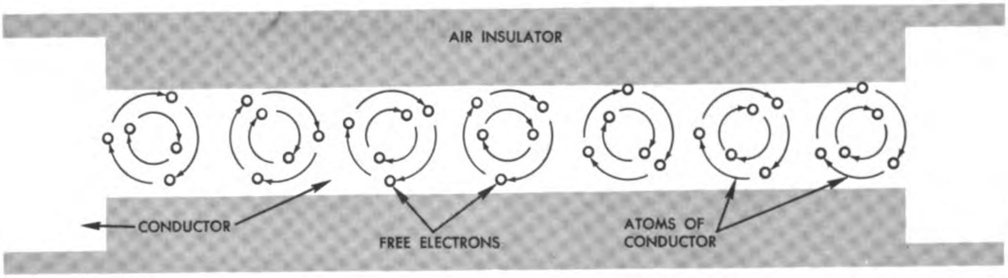
Terms that mean the same as potential difference are electromotive force (EMF), difference of potential (PD), electrical pressure (E), and voltage (V). The term electromotive force is reserved for a battery or other source of the potential difference. The terms voltage, electrical pressure, and difference of potential refer to the potential difference either across a source of energy such as a battery or across a device such as a lamp or a toaster.

ELECTRON FLOW

The electric circuit is defined as a continuous path through which current flows. If an electromotive force (a battery), conductors (wires), and resistances (electric lamps or motors) are combined in such a way as to have current flow, an electric circuit has been made.

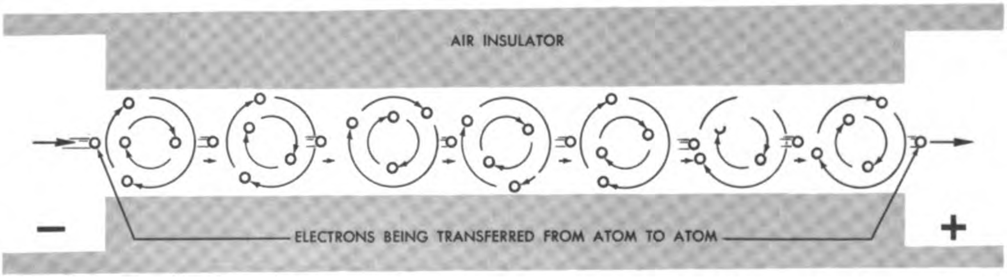
Below is a pictorial explanation of how current flows in a circuit composed of atoms (enlarged many times in the illustration). When free electrons from one atom displace free electrons in a second atom, which in turn displace free electrons in another atom, a continuous motion along a conductor takes place. This electron drift is known as current flow. When an electromotive force or voltage is impressed across a conductor, electrons flow.

A difference of potential determines the direction of electron flow. Electrons flow from the point which has an excess of electrons to a point that has a deficiency of electrons; that is, electrons flow from negative (−) to positive

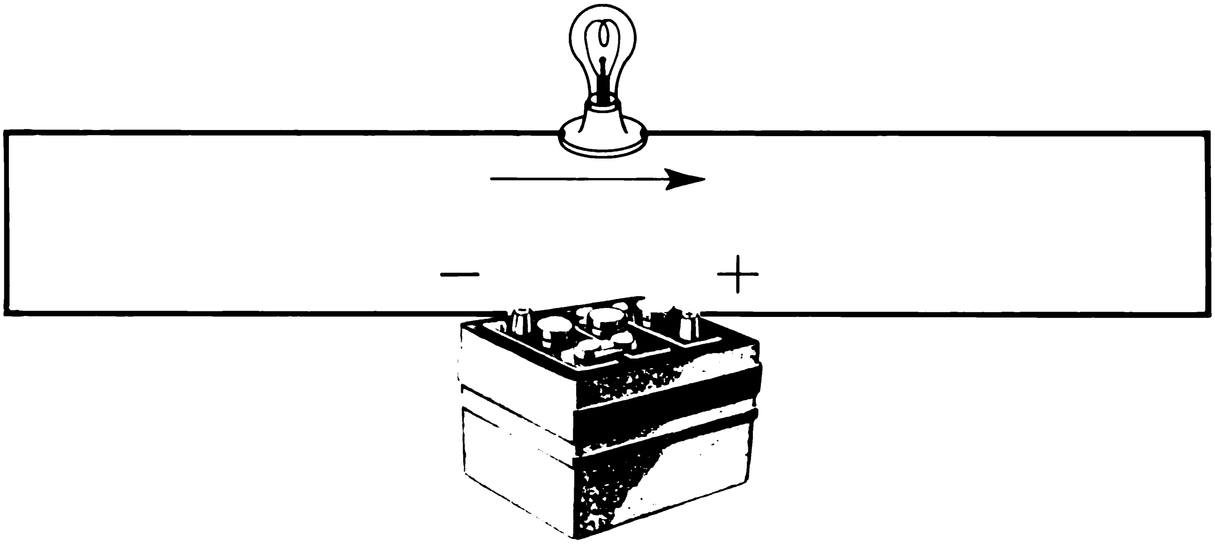


A. CONDUCTOR NORMAL (NO CURRENT)

Movement of Free Electrons in a Conductor



B. CURRENT IN CONDUCTOR



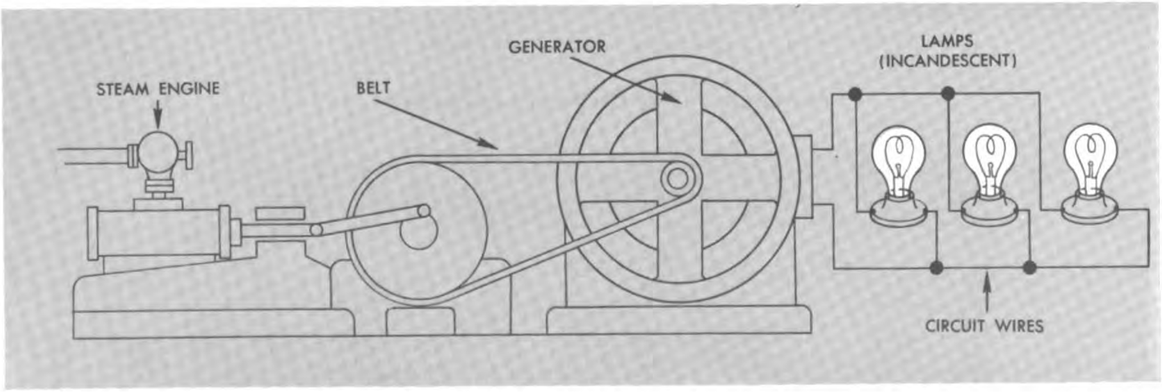
Direction of Current Flow

(+). This concept of negative and positive is referred to as *polarity.*

If the current in the conductor always flows in one direction and is of constant magnitude, it is called *direct current,* or simply DC. If the current is continually varying in magnitude and periodically reversing in direction, it is called *alternating current,* or AC. A current that is always of the same polarity but changes in magnitude at either regular or irregular intervals is called *pulsating direct current.*

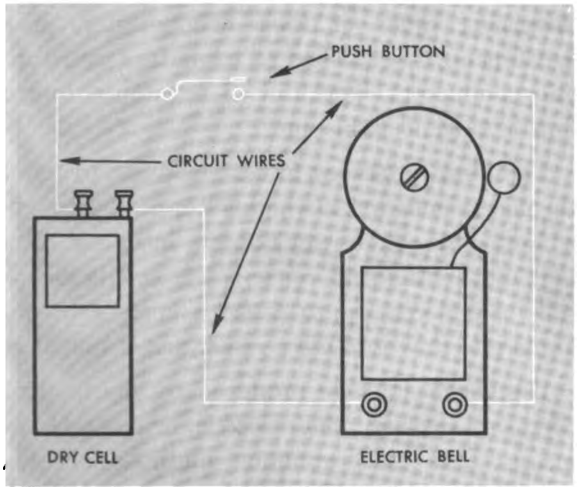
Regardless of whether the current is DC, AC, or pulsating DC, the direction of electron movement *in a conductor* is from negative to positive. *Within a source of electromotive force,* such as a battery, the electrons move from positive to negative. The direction of electron flow both in a conductor and in a source of electromotive force is shown in the illustration at the top of the page.

Electricity, current flow, and many of their effects were discovered long before the electron. The fact that electric current flows in a given direction under a given set of circumstances was also known. It was this fact that gave rise to the concept of polarity. Arbitrarily, a "positive" and "negative" were established, and current was assumed to flow from plus to minus. This concept of plus to minus was borrowed from the laws of mechanics where a body moves from a higher (positive) potential to a lower (negative) potential. As a result, scientific literature was written on the assumption that current flows from plus to minus. After the discovery and



Producing Electron Flow Mechanically

acceptance of the electron, the fallacy of the assumption became obvious. Some authors therefore still make this distinction: plus to minus is current flow and minus to plus is electron flow. In this manual, however, the terms current flow and electron flow are used interchangeably to represent the flow of electrons from minus to plus.



Chemical Production of Electricity

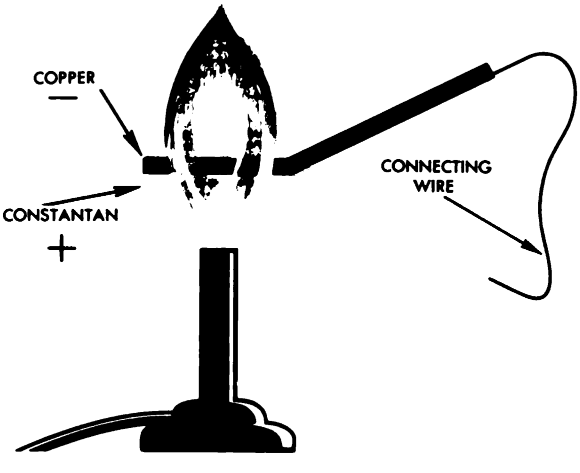
PRODUCING ELECTRON FLOW

A simple method by which electron flow is produced occurs when two bodies with different charges are connected by a conductor. This method serves as an illustration; however, it does not provide current of sufficient magnitude to be of practical use. Other methods have been discovered that are more practicable and that are capable of producing the necessary voltage and current from readily available sources.

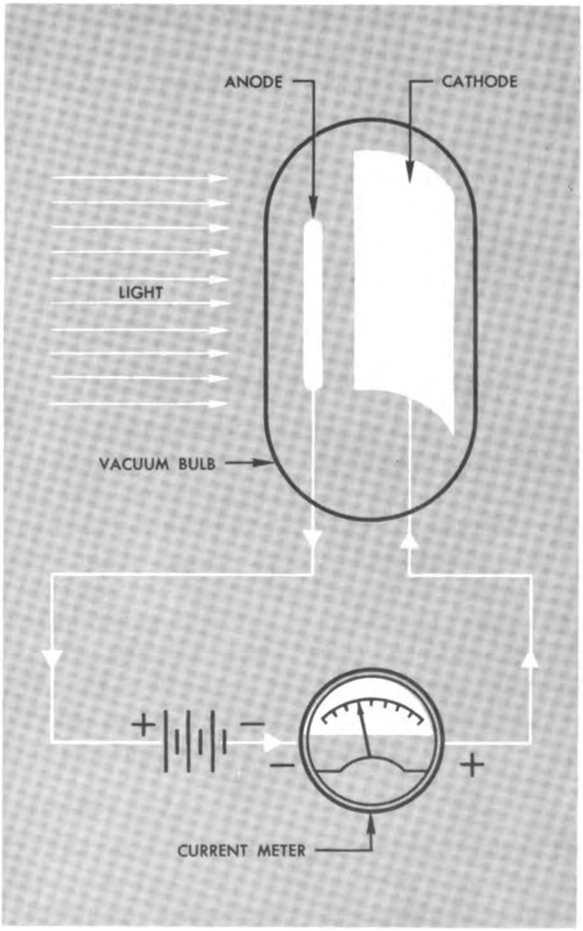
One of the most common means of producing electron flow is *mechanical*. The mechanical energy developed by the steam engine in the illustration on the previous page is transmitted to the electric generator by the belt. The electric generator creates a potential difference which sets the electrons in motion. The moving electrons, or current flow, cause the lights to glow.

The dry cell in the illustration above produces electron flow by *chemical* means.

Current can be produced in a circuit made of two dissimilar metals, if the points of contact are heated, as shown in the illustration above. This method of producing electron flow by heating is called the *thermoelectric* method. The amount of current produced depends upon the



Electron Flow Caused by Heat



A Photoelectric Cell

temperature of the junction. Two dissimilar metals so connected can be used as a temperature-indicating device called a thermocouple.

When light strikes certain surfaces, electrons are emitted; this phenomenon is known as the *photoelectric* effect. A common type of photoelectric cell is shown in the illustration at the bottom to the right on the previous page. If a beam of light is allowed to enter this cell and fall on the sensitive surface, electrons leave this surface. These electrons then flow in the circuit and give an indication on the scale of the current meter. A similar device is commonly used by photographers as an exposure meter.

If a quartz crystal is subjected to mechanical pressure, a potential difference will appear across it. If the crystal is properly connected to a circuit, current will flow. This method of producing electron flow is known as the *piezoelectric* method. The converse of this phenomenon is also true. If a difference of potential is placed across a crystal, it will experience a mechanical stress.

SUMMARY

Current is a flow of free electrons. Its unit is the ampere, which refers to the flow of 6.28×10^{18} electrons (a coulomb) past a point in one second.

Current flows more readily through some substances than it does through others. In the increasing order of their conductivity, such substances are classified as insulators, resistors, and conductors. The unit of resistance to the flow of current is the ohm.

Potential difference forces electrons to move through a substance, whether that substance is a conductor, a resistor, or an insulator. The unit of potential difference is the volt, which is the force necessary to cause a current of one ampere to flow through a resistance of one ohm.

Electricity may be produced by mechanical, chemical, thermal, photoelectric, or piezoelectric means. The most common methods are the mechanical and chemical.

Chapter 4

Introduction to Circuits

An electrical circuit results when a source of electrical potential is connected to an electric device or devices by means of conductors. The electric device may be anything (from a large motor to a small resistor) which requires current for the performance of its function. Since all electric devices have resistance, resistors may be used in circuits to represent these devices in circuits.

CIRCUIT SYMBOLS

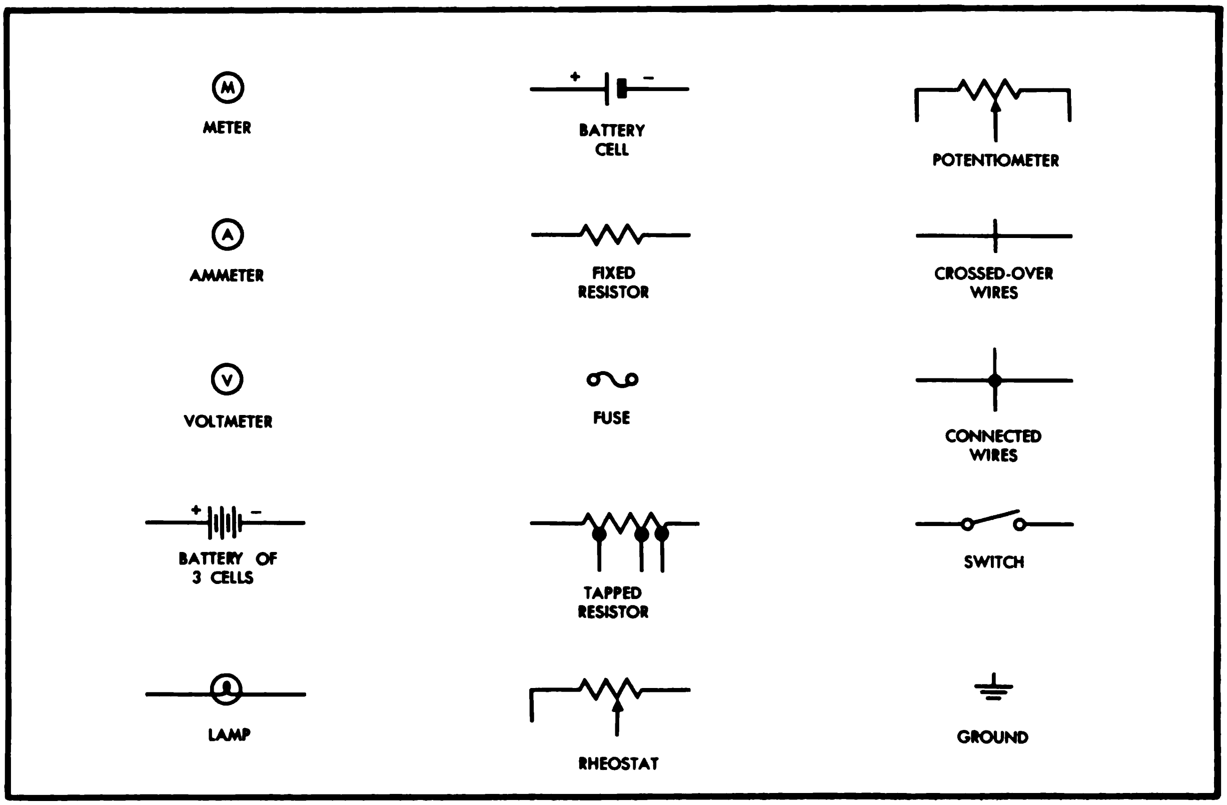
To analyze and understand circuits, you must learn the various symbols that are employed in electrical schematics. A complete chart of circuit symbols is presented in the appendix on pages 553-556. Some symbols which you will need very soon are shown in the illustration on page 18. Learn them so that you can more readily understand the discussions which follow.

CLOSED AND OPEN CIRCUITS

When a circuit is broken so that a current is no longer able to pass, the circuit is said to be an *open circuit.* A complete path of an electric current is a *closed circuit.* A circuit may be closed and made complete by connecting two wires between the source of energy and some electrical device. One wire is for the current to flow to the electrical device from the power supply. The other wire is to allow the current to return to the source of energy. Sometimes a circuit is made complete by *grounding* it.

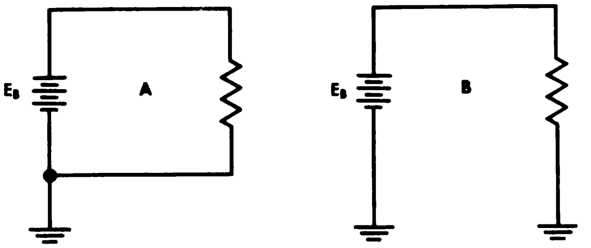
Ground

The term *ground,* which actually means the earth, is often used to describe a zero reference for voltage measurements and a point of common return from one side of circuit components to the same side of the power source. In the



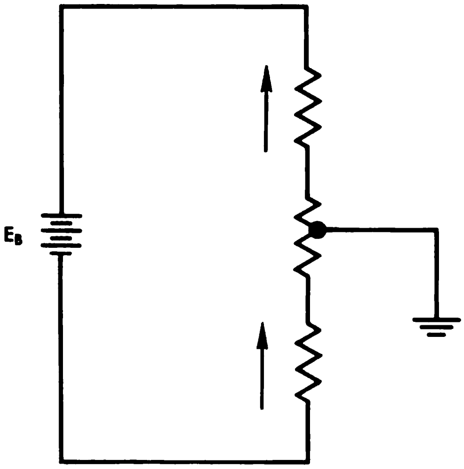
Some Standard Electrical Symbols

illustration below, circuits A and B are complete circuits that are identical electrically, even though the two grounded points in B are separated physically. The arrangement in B is common in practical circuits, as it eliminates a great number of wires. An automobile is an example. One side of the battery is grounded to the frame; so, too, is one side of each electrical unit — lights, generator, spark plugs, and starting motor. The return path for all of these circuits is through the frame of the automobile. The same sort of arrangement is used in other circuits.



Complete Circuits

It makes no difference whether the negative or positive terminal of a battery is grounded, provided the same terminal is grounded in all cases. In the circuit shown below, ground is being used as a point of zero reference for voltage measurements. In this case neither the



Ground Is a Zero Reference

positive nor the negative side of the battery is grounded. Arrows show the direction of current flow in each part of the circuit. Current flow is from the negative side of the power source to the positive side.

OHM'S LAW

You have seen that an electrical circuit consists of a source of potential connected to an electrical device which has resistance. When the circuit is completed, current flows in a closed path.

In 1827 George Simon Ohm discovered that the amount of current which flows in a circuit increases if the potential applied is increased, and decreases if the resistance of the circuit is increased. He also found that the current decreased when he decreased the potential applied, and increased if the resistance were decreased. These findings are known as Ohm's law which states that *the current in a circuit is directly proportional to the applied potential and inversely proportional to the resistance.* It is written mathematically as follows:

$$I = \frac{E}{R}$$

where I is the current in amperes,
E is the EMF in volts,
R is the resistance in ohms.

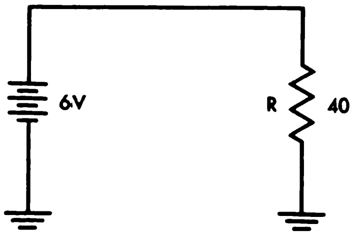
The formula may be solved for E in terms of I and R, and for R in terms of E and I. The results are as follows:

$$E = I \times R$$

$$R = \frac{E}{I}$$

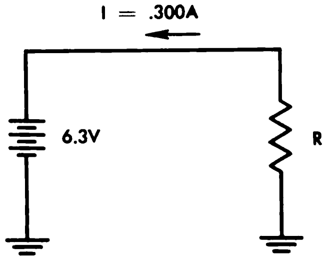
Ohm's law may be used to solve various problems as shown below.

1. A lamp of 40-ohms resistance is connected to a battery of 6 volts. How much current does the lamp draw?



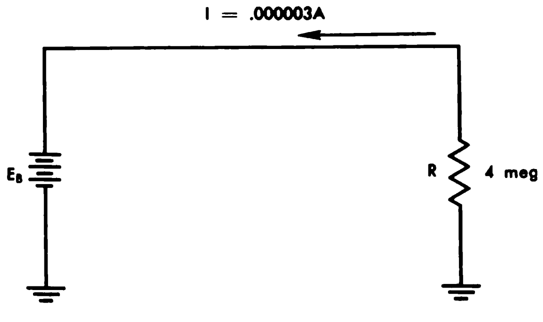
Given: $E = 6$ volts Find: I
$R = 40$ ohms
Solution: $I = \frac{E}{R}$
$I = \frac{6}{40}$
$I = .15$ amps

2. What is the resistance of the filament of a light bulb which requires 6.3 volts and draws 300 ma.



Given: $E = 6.3$ volts Find: R
$I = 300$ ma.
Solution: 300 ma. = .3 amps
$R = \frac{E}{I}$
$R = \frac{6.3}{.3}$
$R = 21$ ohms

3. A 4-megohm resistor has 3 microamperes of current flowing through it. What is the voltage across the resistor?



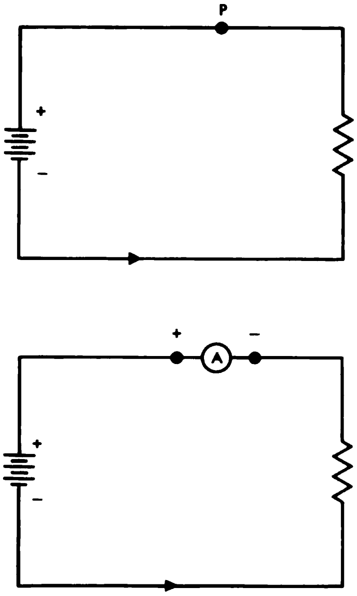
Given: $R = 4$ megohms Find: E
$I = 3$ microamperes
Solution: 4 megohms $= 4 \times 10^6$ ohms
3 microamperes $= 3 \times 10^{-6}$ amps
$E = I \times R$
$E = 4 \times 10^6 \times 3 \times 10^{-6}$
$E = 12$ volts

MEASURING DEVICES

The measurement of electrical quantities is a very important part of electronics. Electrical indicating instruments have been devised for measuring the various electrical quantities. Such instruments are tools, and you need to master their use.

Ammeter

The electrical instrument used to measure the amount of current in an electrical circuit is called an *ammeter*. It is connected in such a manner that it will measure all the current passing by a given point P. The illustration below shows how an ammeter is hooked into the circuit to measure the current passing a point. When a meter is connected so all of the current flows through it, it is said to be in *series*. It is necessary to observe *polarity* in connecting an ammeter.

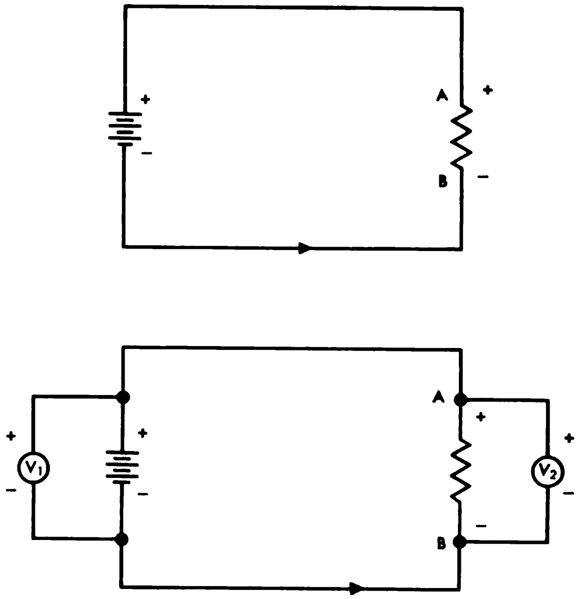


Connecting an Ammeter

An ammeter that measures current in milliamperes is called a milliammeter. One that measures current in microamperes is a microammeter.

Voltmeter

The electrical instrument used to measure difference of potential, or voltage, is called the *voltmeter*. It is placed in such a manner as to measure voltage between the two points A and B. The illustration above shows the proper connection of a voltmeter. To measure the emf of the battery, connect voltmeter V_1 as shown, observing polarity. To measure the potential difference between points A and B, it is necessary to trace the electron flow and observe the polarity between these points. Note that the side of the resistor that the electrons enter is the negative side. A meter that is hooked across a difference of potential is said to be hooked in *parallel*.



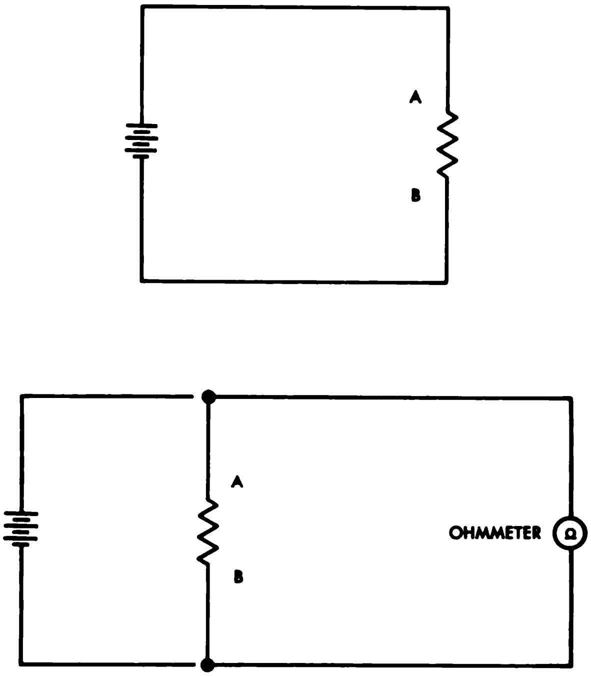
Connecting a Voltmeter

Ohmmeter

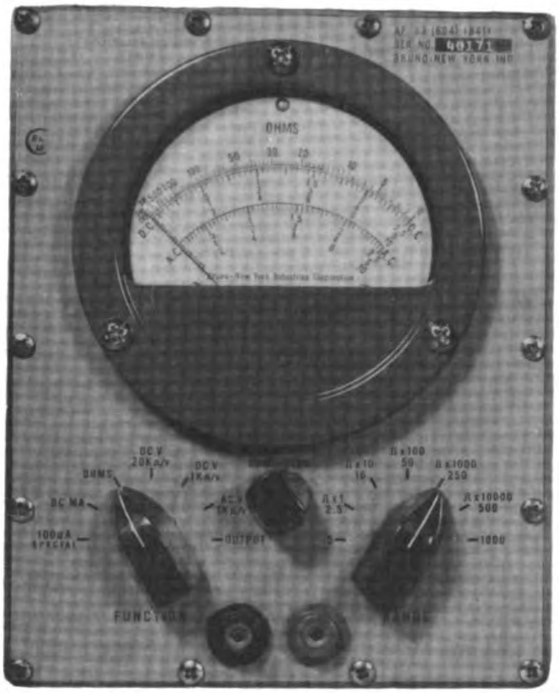
The ohmmeter is an electrical instrument which is used to measure the amount of resistance in a circuit. The resistance of a circuit is one of the factors which determines its proper operation. Faults in circuits often can be located by checking the values of their resistances.

The proper method of measuring a resistance with an ohmmeter is illustrated on the next page. Note that the resistance AB is disconnected from the circuit before placing the meter across it.

A closed circuit provides an external pathway for electrons to move from the negative terminal of a voltage source to the positive



Measuring Resistance with an Ohmmeter



Multimeter IS-189

terminal. A closed circuit always has some value of resistance. It may be low or high depending upon the circuit. An ohmmeter may therefore be used to check a closed circuit. An indication of resistance means that the circuit is closed. This is known as checking for *continuity*.

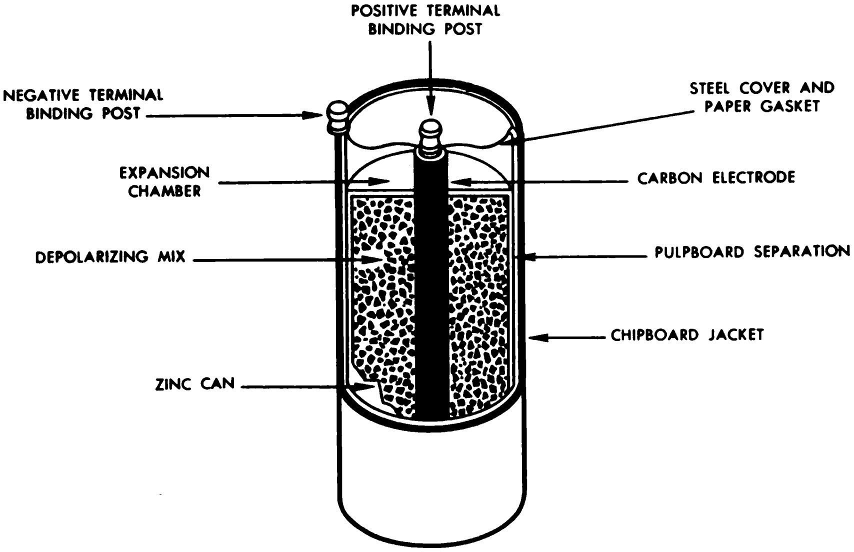
If an ohmmeter indicates a zero reading between two points in a circuit, there is no resistance between them. This means that a *short circuit* exists between the points, and electron flow is unopposed.

If an ohmmeter indicates a reading of infinity between two points, there is infinite resistance between them. This condition means that an *open circuit* exists between the points, and electrons cannot flow.

Each of the meters discussed consists of a basic meter movement combined with additional devices necessary for its particular purpose. A voltmeter, ohmmeter, and ammeter may be combined in one unit using only one meter movement. Switches, jacks, and the necessary additional devices are included to make such a combination volt-ohm-milliameter or *multimeter*. By proper arrangement of parts, the multimeter can be built into a small, compact unit. The illustration above shows a typical multimeter often used by observers to check equipment.

THE DRY CELL

As you have already seen, the dry cell produces electron flow by chemical means. A very common use of this source of EMF is the ordinary flashlight cell. The construction of a dry cell is shown in the illustration on page 22. The positive terminal, or electrode, of the dry cell is a fairly large carbon rod located in the center of the zinc container that forms the negative electrode. The electrolyte, which is a dilute solution of ammonium chloride (sal ammoniac), is mixed with some porous, inert material to form a paste which is placed on the inside of the zinc container. Between this paste and the carbon rod, located in the center, is a porous mass of manganese dioxide and carbon granules called a *depolarizer*. Its function is to absorb the ammonia and hydrogen gases that are given off. *Polarization*, an accumulation of hydrogen gas at the positive electrode, would have the detrimental effect of insulating the electrode.

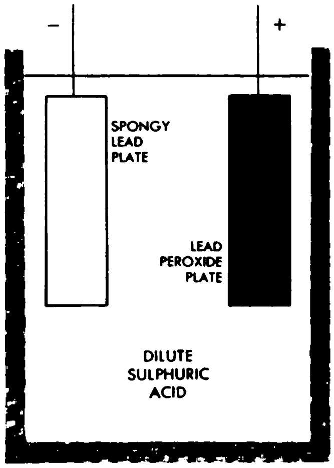


A Typical Dry Cell

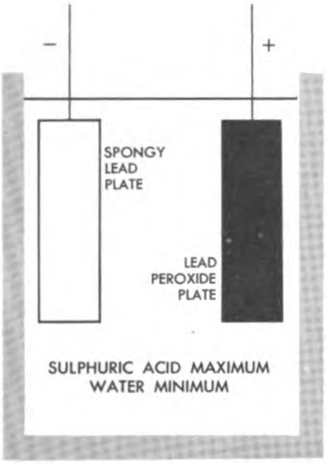
The top of the cell is sealed by an insulating compound so that only the terminals connected to the carbon rod and zinc container are visible. The zinc container is then placed in a cardboard carton, which acts as a protector and an insulator for the dry cell. Each dry cell has an electromotive force of 1.5 volts. A defect manifested in some cells when impurities in the zinc cause the zinc to be eaten away too rapidly is known as *local action.* It may be offset by a process known as amalgamation, which prevents the electrolyte from acting on the impurities in the zinc.

THE STORAGE BATTERY

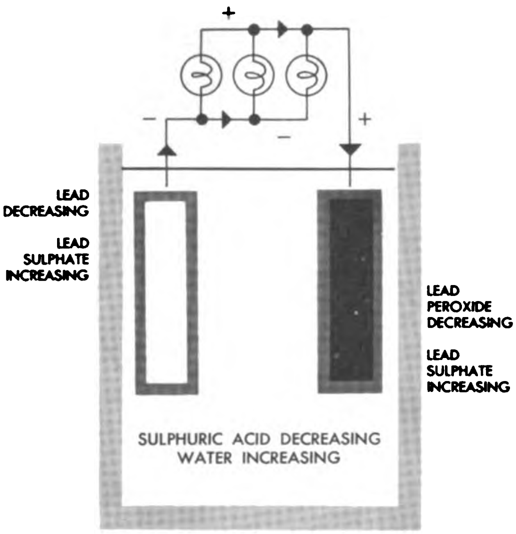
The storage battery is another method of producing electron flow by chemical means. The lead-acid cell shown in the illustration to the right is the most common type. It consists of a positive plate of lead peroxide and a negative plate of porous or spongy lead in a dilute solution of sulphuric acid. The usual construction of such cells of the portable type is known as a posted-grid construction. A lead and antimony alloy is used to make the grid. This grid is filled with a paste which, after a forming process, produces lead peroxide as the active chemical material. The grid acts as physical support for the chemical and as an electrical conductor for the current that is developed by chemical action. A similar grid is used for the negative plate; but a different paste is used which, after the forming process, consists mainly of pure lead in a very porous condition. These plates are supported in a rubber or glass jar containing the electrolyte, composed of dilute sulphuric acid. The chemical reactions that take place are rather complicated, but the final result is that both plates change their active material to lead sulphate as the cell is discharged. This lead sulphate has a tendency to harden, crystallize, and expand if permitted to stand. This action is known as *sulphation.* For this reason, it is important to keep lead cells in a charged condition, with sufficient water added to make up for evaporation.



A Lead Acid Cell



CHARGED



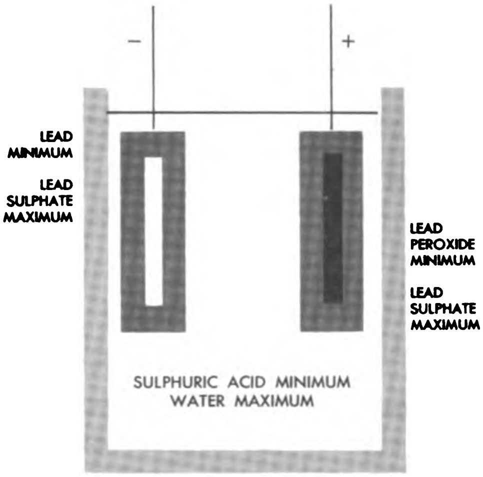
DISCHARGING

Charge and Discharge of a Storage Cell

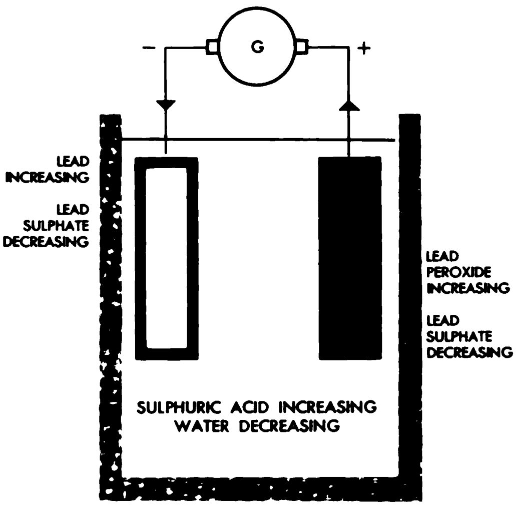
The illustration on this page shows the chemical action in a lead-acid cell or storage battery. In the charged state, the sulphuric acid in the electrolyte is at a maximum and the water, at a minimum; neither plate has lead sulphate. While the battery is discharging, the sulphuric acid in the electrolyte is decreasing and the water is increasing; both plates are changing to lead sulphate. In the discharged state, the sulphuric acid is at a minimum and the water, at a maximum; both plates have a maximum of lead sulphate. To recharge the battery a DC generator is connected which changes conditions back to the charged state by applying an electric current.

Each cell has an electromotive force of about 2.1 volts. Two or more lead-acid cells are ordinarily used in a combination to form a storage battery. This combination provides increased voltage output, increased current output, or both. The automobile battery is a 6-volt battery consisting of three lead-acid cells.

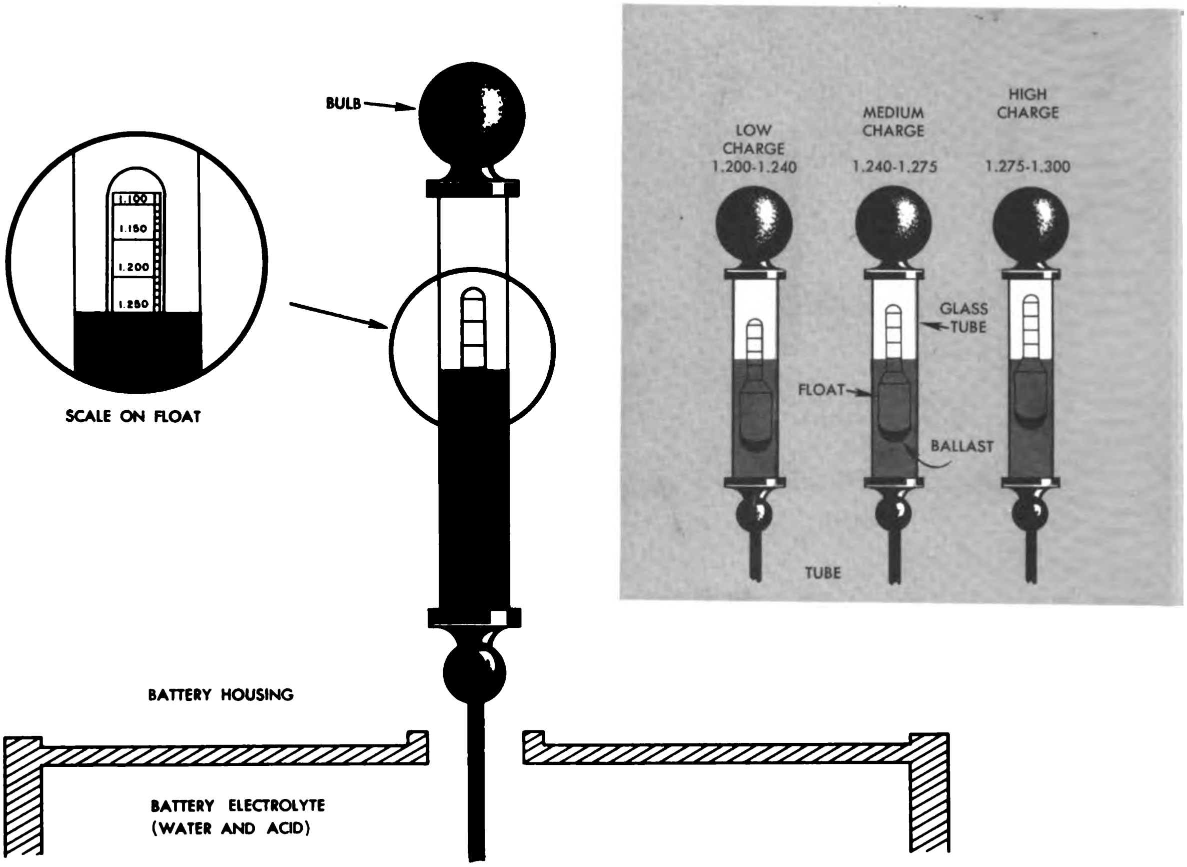
The *hydrometer* is a device used to check the state of charge of a storage cell by the concentration of the electrolyte. The use of a hydrometer to check the charge of a lead-acid storage battery is shown in the illustration on page 24.



DISCHARGED



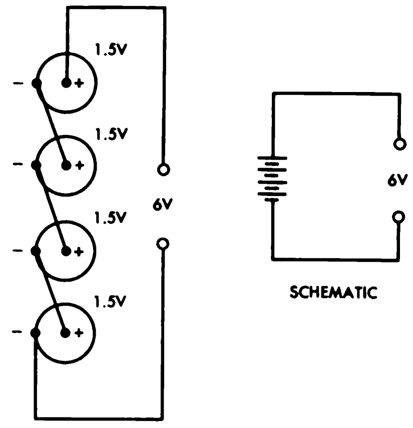
CHARGING



Using a Hydrometer

BATTERY CONNECTIONS

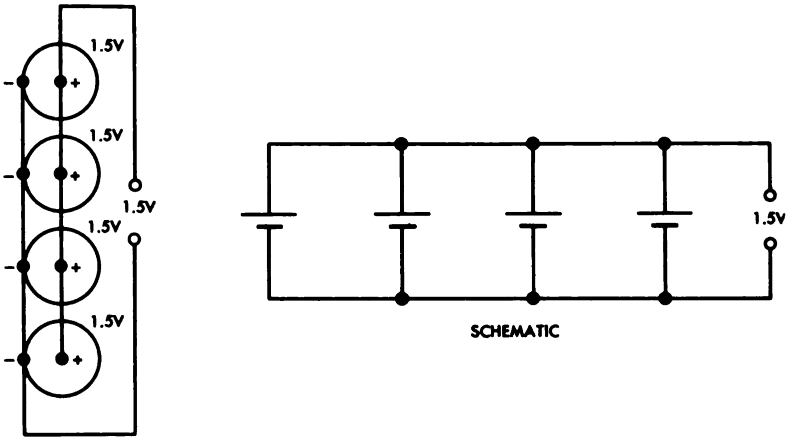
Some dry cells such as the flashlight cell are used individually as a source of electric potential for an electrical circuit. In other cases, several dry cells are connected so as to increase the voltage, the current, or both. Such a combination is a BATTERY and is typified in the well-known B-battery used in portable radio equipment.



Cells in Series

When the positive terminal of one cell is connected to the negative terminal of the next cell, the cells are said to be connected in *series*. Under these conditions, the voltage appearing at the two ends of the series of cells is the sum of the individual cell voltages. If a voltmeter is placed across the two free ends of the cells in series, in the illustration to the left, the voltmeter would read a total of 6 volts. An example of such a connection of cells is a 4-cell flashlight.

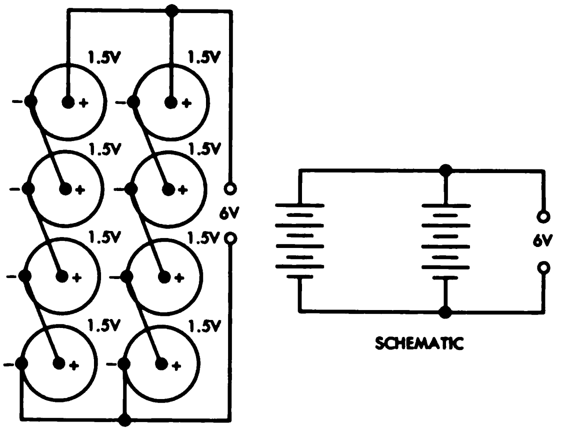
When the positive terminal of one cell is connected to the positive terminals of other cells, and the negative terminals are tied together, as shown in the illustration on the next page, the cells are said to be connected in *parallel*. Under these conditions the terminal



Cells in Parallel

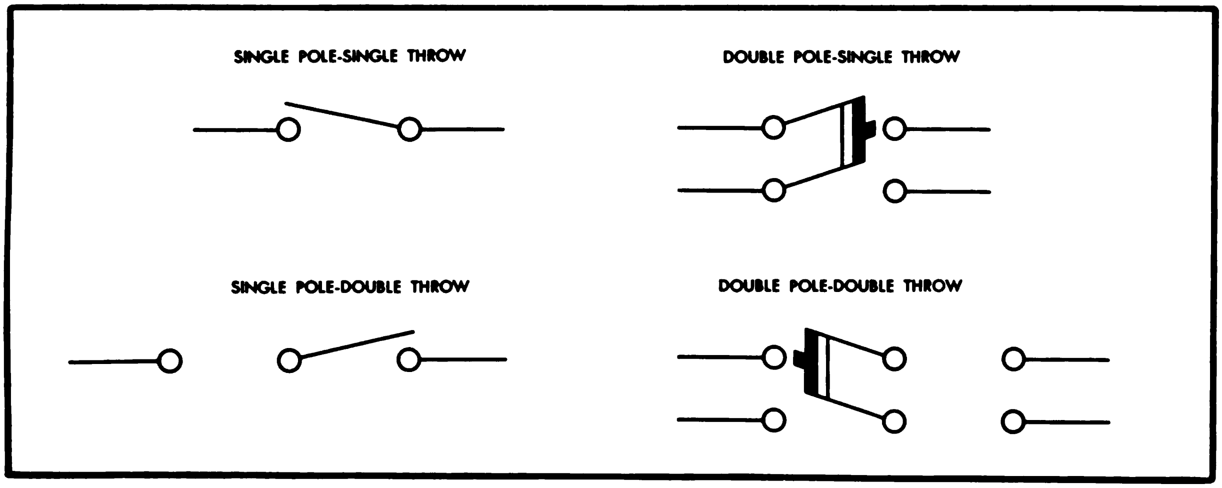
voltage of the combination is the same as the terminal voltage of each cell. A circuit whose power requirements are low voltage and heavy current drain requires cells connected in parallel.

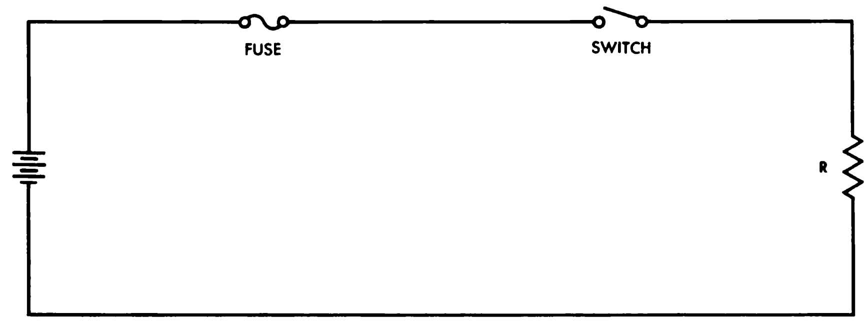
Cells may also be connected in what is called a series-parallel arrangement. In the illustration, two sets of cells are in series, and the sets themselves are connected in parallel. The terminal voltage is the same as the voltage appearing across any one set of cells in series. A circuit whose power requirements are high voltage and heavy current drain requires such an arrangement of cells.



Cells in Series Parallel

Typical Switches





Circuit with Fuse and Switch

SWITCHES AND FUSES

A *switch* is a device for making or changing connections in an electrical circuit. There are various types of switches: single pole-single throw, double pole-single throw, single pole-double throw, double pole-double throw, push button, etc. The illustration on the previous page shows the schematic symbols of some typical switches found in electronic equipment.

A *fuse* is a device used to protect electrical equipment from being damaged because of excessive current flow. It works on the principle that current flowing through a resistance causes heat. The greater the current, the greater is the heat produced. As the current which flows through the low resistance of the fuse becomes excessive, the heat becomes sufficient to melt the fuse and open the circuit. This prevents further current flow and protects the circuit from damage. Fuses are of many different types and range in size from a few milliamperes to hundreds of amperes. Since fuses are used to protect equipment from overload, do not replace a fuse by one of a larger rating. If you do, you are defeating the purpose for which the fuse is intended.

The illustration above shows how a fuse and a switch are represented schematically in a circuit. The resistor R represents any device which draws current.

Chapter 5

Series Circuits

A series circuit is one in which resistances or other electrical devices are connected end to end as shown in the illustration on page 28. Note that the path of current flow is from the negative side of the battery through resistors R_1, R_2, and R_3 and back to the positive side of the battery. A series circuit may therefore be defined as a circuit in which *the current flows in only one path*. The current which flows is the same in each part of the circuit. Expressed mathematically,

$$I_t = I_1 = I_2 = I_3$$

where I_t is the total current and I_1, I_2, and I_3, the current through R_1, R_2, and R_3, respectively.

TOTAL RESISTANCE

The current which flows in the series circuit meets opposition from each resistor in turn. If the resistors R_1, R_2, and R_3 were 30, 40, and 50 ohms, respectively, the current in the circuit would meet a total opposition of 120 ohms. The three resistors would then offer the same opposition to the current in the circuit as one resistor of 120 ohms. In a series circuit, *the total resistance is equal to the sum of the individual resistances*. Expressed mathematically,

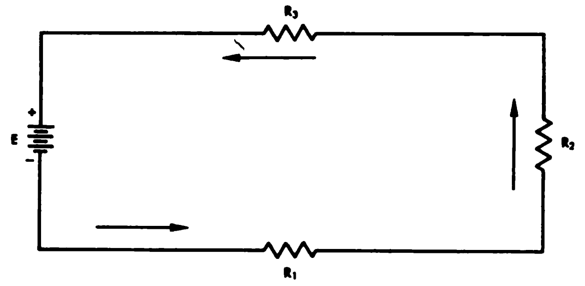
$$R_t = R_1 + R_2 + R_3$$

where R_t is the total resistance and R_1, R_2, and R_3 are the resistances in series.

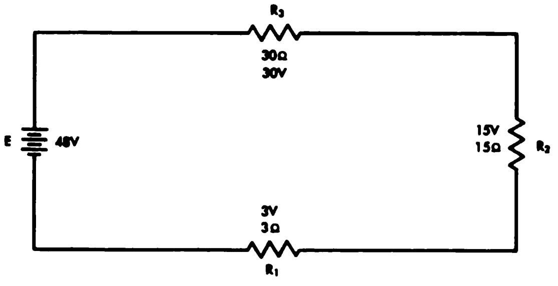
VOLTAGE DROPS

As the current flows in a series circuit, a voltage is developed across each resistor. The voltage developed across each resistor can be found by Ohm's law if the values of the current and the resistor are known.

In the series circuit, a voltage of 48 volts is applied across resistors R_1, R_2, and R_3. Since the total resistance is equal to the sum of the three resistances, $R_t = 48$ ohms. By Ohm's



A Series Circuit



Voltage Drops in a Series Circuit

law, $I = \frac{E}{R}$, the current in the circuit is found to be 1 amp. The voltage across each resistor can be found by Ohm's law, $E = IR$. Thus the *voltage drop* across R_1 is 3 volts, across R_2 is 15 volts, and across R_3 is 30 volts.

Note that the sum of the voltage drops E_1, E_2, and E_3 across R_1, R_2, and R_3, respectively, is equal to the applied voltage of 48 volts. In series circuits, ***the sum of the voltage drops across the resistors is equal to the applied voltage.*** Expressed mathematically,

$$E_t = E_1 + E_2 + E_3$$

where E_t is the applied voltage and E_1, E_2, and E_3 are the voltage drops across R_1, R_2, and R_3, respectively.

POWER

Before discussing power in electrical circuits, you will find it of value to learn the meanings of the following terms commonly used in electronics: force, work, energy, and power.

Force is defined as that which produces, or tends to produce, a change in motion of a body. In electronics, force causes electrons to move from one point to another. The unit of force is the *dyne.*

Work is the production of motion against a resisting force. In electronics, work is done when electromotive force causes electrons to move against the opposing force offered by the resistance in a circuit. When an ampere of current flows through a resistance of one ohm for one second, a *joule* of work is done.

Energy is the capacity or ability to do work. Energy that is due to motion of matter is called *kinetic* energy, while energy that is due to the position of matter is called *potential* energy. As an example, the electrons on a negatively charged body possess potential energy with respect to a less negatively charged body, because these electrons would flow if the two bodies were connected by a conductor. The electrons moving along the conductor connecting the two bodies possess kinetic energy because such electrons are in motion. Energy may be manifested in such forms as heat, light, and motion. It may be transformed from one form to the other as is done in producing electron flow by mechanical, chemical, thermal, photoelectric, or piezoelectric means. Since energy is the capacity to do work, energy and work have the same units.

Power is the rate of doing work and is abbreviated P. Forcing electrons through a resistor requires work, and energy is expended in the resistor in the form of heat. If this heat is not too great, it can be radiated and will not damage the resistor. If it is excessive, it cannot be radiated as quickly as it is generated; consequently, the resistor becomes too hot and burns. As a precaution against excessive heating, a resistor is specified as having a certain power rating, meaning that no more than a specified current (given number of electrons per second) can be allowed to flow through the resistor without damaging it.

The unit of electrical power is the *watt*. One watt of power is present in a one-ohm resistor in which a current of one ampere is flowing, in other words, a resistor consuming one joule of energy per second. A larger unit is the *kilowatt*, which equals 1000 watts. The power dissipated by most of the resistors used in the laboratory is of the order of one-tenth of a watt. However, some resistors dissipate as much as fifty watts. A *horsepower*, the unit for measuring mechanical power, is equivalent to 746 watts.

In a DC circuit, power is equal to the voltage times the current. Expressed mathematically,

$$P = E \times I$$

where P is in watts, E is in volts, and I is in amperes.

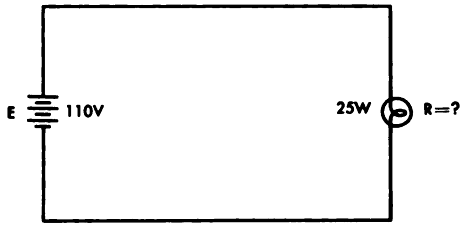
Using Ohm's law, you can derive two other equations for power in a DC circuit. Thus, substituting IR for E you obtain the equation $P = I^2R$, and substituting $\frac{E}{R}$ for I, you obtain the equation $P = \frac{E^2}{R}$.

Since power is the rate of doing work or consuming energy, the length of time power is being used is the true measure of energy consumed. It is common practice to purchase electrical energy by *watt hours* (watts×hours). This unit is so small that the unit *kilowatt-hours*, equal to one thousand watt-hours, is used. As an example, a 100-watt lamp requires 100 watts of power for proper operation and consumes 100 watt-hours of energy in 1 hour. In terms of kilowatt-hours, the lamp uses $\frac{100}{1000} = 0.1$ kilowatt-hour in 1 hour. In ten hours of operation, the lamp would consume 1 kilowatt hour of energy. In twenty-four hours, the lamp would consume 2.4 kilowatt-hours. At five cents per kilowatt hour, this would cost twelve cents.

SERIES CIRCUIT PROBLEMS

The formulas for power and Ohm's law may be used to solve many series circuit problems. The following illustrative problems show how these formulas are used:

1. What is the resistance of a 25-watt lamp designed to operate on 110 volts?



Given: $E = 110$ volts Find: R

$P = 25$ watts

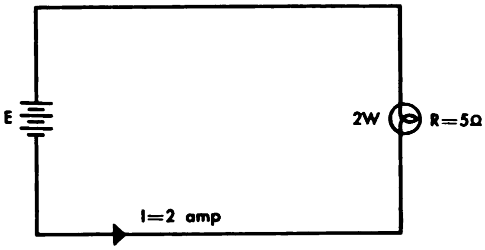
Solution: $P = \frac{E^2}{R}$

$$25 = \frac{110 \times 110}{R}$$

$$R = \frac{12{,}100}{25}$$

$$R = 484 \text{ ohms}$$

2. What will happen to a 2-watt lamp whose resistance is 5 ohms if the current flow through it is 2 amperes?



Given: $I = 2$ amp. Find: P

$R = 5$ ohms

Solution: $P = I^2R$

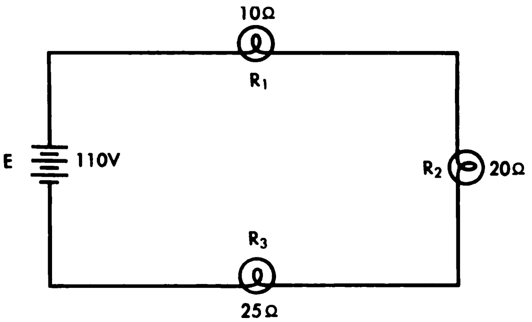
$$P = 2^2 \times 5$$

$$P = 20 \text{ watts}$$

The lamp will burn out very soon because its power rating is being exceeded.

3. Three lamps are connected in series across a 110 volt line. If they have resistances of 10, 20, and 25 ohms, respectively, find: the total resistance, the current flow, the voltage drop

across the lamp L_2, the power expended in the circuit, and the number of kilowatt-hours consumed.



Given: $E = 110$ volts Find: R_t, I, E_2, P, KWH

$R_1 = 10$ ohms, $R_2 = 20$ ohms, $R_3 = 25$ ohms

Time = 12 hours

Solution: $R_t = R_1 + R_2 + R_3$

$R_t = 10 + 20 + 25$

$R_t = 55$ ohms

$$I = \frac{E_t}{R_t}$$

$$I = \frac{110}{55}$$

$I = 2$ amps

$E_2 = IR_2$

$E_2 = (2)\ (20)$

$E_2 = 40$ volts

$$P = EI \qquad P = I^2R \qquad P = \frac{E^2}{R}$$

$$P = (110)\ (2) \qquad P = 2^2\ (55) \qquad P = \frac{110 \times 110}{55}$$

$$P = 220 \text{ watts} \qquad P = 220 \text{ watts} \qquad P = 220 \text{ watts}$$

$$KWH = \frac{220 \times 12}{1000}$$

$KWH = 2.64$

KIRCHHOFF'S LAWS

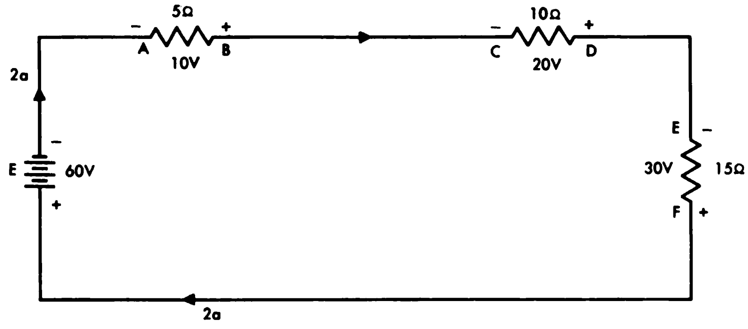
In 1847 the German physicist, Gustav Richard Kirchhoff, elaborated on Ohm's law and developed what is known as Kirchhoff's laws for current and voltage. These laws may be stated as follows:

The algebraic sum of the currents at any junction of conductors is zero. This means that at any point in a circuit there is as much current flowing toward the point as there is flowing away from it.

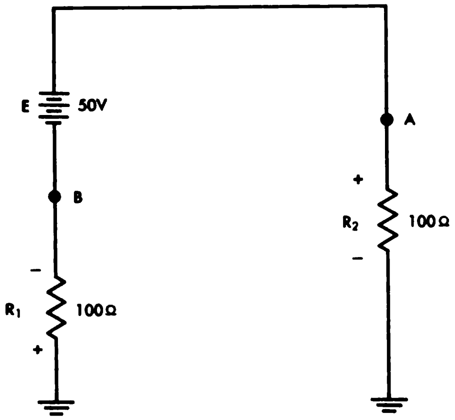
The algebraic sum of the EMFs and voltage drops around any closed circuit is zero. This means that in any closed circuit the applied EMF is equal to the sum of the voltage drops around the circuit.

Polarity

In the statement of Kirchhoff's laws the term *algebraic sum* was used. An algebraic sum differs from an arithmetic sum in that both the magnitude and the sign of each number must be considered. In electronic circuits, a voltage drop occurs when current flows through a resistor. The magnitude of the voltage is determined by the size of the resistor and the amount of current flow. However, the direction of the current flow determines the *polarity* (sign) of the voltage drop. For example observe the polarities of the applied EMF and the voltage drops as indicated in the illustration below. The applied EMF causes electrons to flow through the opposition offered by the resistances. The voltage drop across each resistance is therefore opposite in polarity to that of the applied EMF as shown. Trace the electron flow from the negative side of the battery through the resistors and back to the positive side of the battery. Note that the side of each resistor through which the current enters is labeled the negative side.



Polarity of Voltage Drops



Positive and Negative Voltage Drops

A similar situation exists when ground is used as a point of reference. If you know the direction of the current flow in a circuit, you may determine the polarity of the voltage across each resistor. In the illustration shown above, point A is positive with respect to ground, while point B is negative with respect to ground. By Ohm's law you can compute the voltage across each resistor.

$$I = \frac{50}{200} = 0.25 \text{ amperes}$$

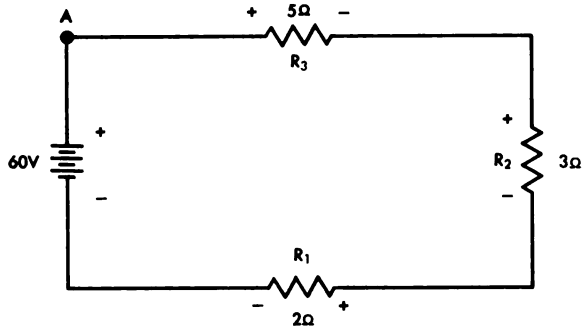
$$E = .25 \times 100$$

$$E = 25 \text{ volts}$$

Measurement with a voltmeter would agree with your answer. However, in using a voltmeter, you must observe correct polarity. By placing one lead of the voltmeter on the reference point (ground) and the other on point A, you measure 25 volts positive with respect to ground. Using ground and point B, you would measure 25 volts negative with respect to ground. Note that a negative voltage is just as effective as a positive one, and the same safety precautions must be exercised with each.

Application of Kirchhoff's Second Law

Kirchhoff's second law (for voltage) may be applied in the circuit shown in the illustration at the top of the next column. The total resistance is the sum of R_1, R_2, and R_3 which is 10 ohms. Since the applied EMF is 60 volts, the current flowing in the circuit is 6 amps. Therefore, the voltage drops across R_1, R_2, and R_3 and are 12 volts, 18 volts, and 30 volts, respectively. The sum of these voltage drops is equal to the applied voltage of 60 volts.



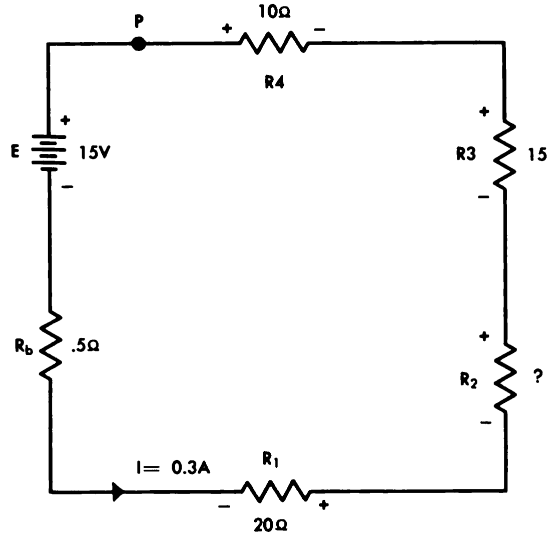
Applying Kirchhoff's Law for Voltage

You may also solve this circuit by using the polarities of the voltages and showing that the algebraic sum of the voltages is zero. Consider voltages positive if the + sign is met first and negative if the − sign is met first as you trace electron flow. By starting at A and going in the direction of electron flow, you may set up the following equation:

$$E_t = +60 - 12 - 18 - 30$$

$$E_t = 0$$

The point at which to start around the circuit and the choice of which polarity is to be considered positive is arbitrarily selected and is a matter of choice for each circuit.



Solving for an Unknown Resistor

Kirchhoff's second law may also be used for finding an unknown resistor. The circuit shown in the illustration at the bottom of the preceding page has three known resistors and one unknown resistor connected across a 15-volt battery whose internal resistance is 0.5 ohm. If a current of 0.3 amp flows in the circuit, what is the value of the unknown resistor?

Given: Find: R_2

$E_b = 15$ volts

$R_1 = 20$ ohms, $R_3 = 15$ ohms

$R_4 = 10$ ohms $R_b = 0.5$ ohm

$I_t = 0.3$ amp

Solution:

$15 = .15 + 3 + 45 + .3R_2 + 6$

$.3R_2 = 1.35$

$R_2 = 4.5$ ohms

The same result may be obtained by starting around the circuit from point P and observing polarity. You obtain the following equation:

$$+15 - .15 - 3 - 4.5 - 3.R_2 - 6 = 0$$

$$-3R_2 = -1.35$$

$$R_2 = 4.5 \text{ ohms}$$

SUMMARY OF SERIES CIRCUITS

The current in a series circuit is the same at all points in a circuit.

The total applied voltage of a series circuit is equal to the sum of the voltage drops across the components.

The total resistance of a series circuit is equal to the sum of the individual resistances.

The total power dissipated by a series circuit is equal to the sum of the power dissipated by the components.

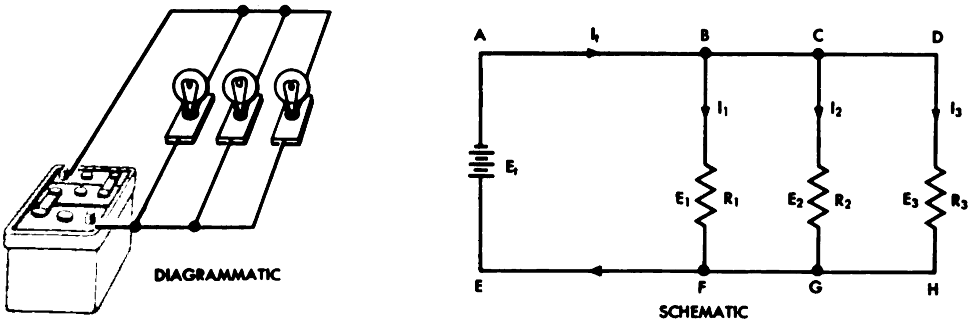
Chapter 6

Parallel Circuits

It is often necessary to connect electrical devices so that the entire source of voltage is across each device. A circuit in which two or more devices are connected across the same source of EMF is a *parallel* circuit. The illustration below is a parallel circuit. It consists of a battery, and three electric lamps connected in parallel. Note that the current which leaves one terminal of the battery breaks up into three parts and then recombines as it returns to the other terminal of the battery. Parallel circuits have ***more than one current path.***

VOLTAGE AND CURRENTS

In the schematic diagram of the parallel circuit shown, note that points A, B, C, and D are connected together and are one point electrically. Similarly points E, F, G, and H comprise another electrical point. Since the applied voltage appears between points A and E, the same voltage appears between points B and F, between points C and G, as well as between points D and H. The definition of a parallel circuit may, therefore, be expressed mathematically as follows:



A Parallel Circuit

$$E_t = E_1 = E_2 = E_3$$

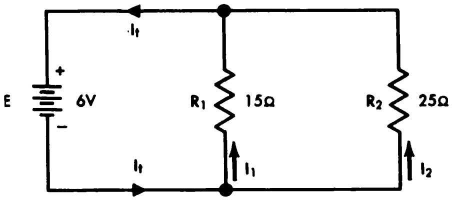
Where E_t is the applied voltage, E_1 is the voltage across R_1, E_2 across R_2, E_3 across R_3.

Kirchhoff's first law states that the current flowing toward a point is equal to the current flowing away from the point. Therefore,

$$I_t = I_1 + I_2 + I_3$$

where I_t is the total current and I_1, I_2, and I_3 are the currents through R_1, R_2, and R_3, respectively.

To find the total current flowing in the parallel circuit shown below, apply the above equations.



Finding Total Current

Given: $E = 6$ volts Find: I_t

$R_1 = 15$ ohms

$R_2 = 25$ ohms

Solution:

$$I_1 = \frac{E}{R_1} = \frac{6}{15} = .4 \text{ amp}$$

$$I_2 = \frac{E}{R_2} = \frac{6}{25} = .24 \text{ amp}$$

$$I_t = I_1 + I_2$$

$$I_t = .4 + .24 = .64 \text{ amp}$$

TOTAL RESISTANCE

In a parallel circuit you have seen that $I_t = I_1 + I_2 + I_3$. By Ohm's law, you may obtain the following relationships: $I_t = \frac{E_t}{R_t}$, $I_1 = \frac{E_1}{R_1}$, $I_2 = \frac{E_2}{R_2}$, and $I_3 = \frac{E_3}{R_3}$.

Substitute these values in the equation for total current and you have, $\frac{E_t}{R_t} = \frac{E_1}{R_1} + \frac{E_2}{R_2} + \frac{E_3}{R_3}$

But in a parallel circuit $E_t = E_1 = E_2 = E_3$.

Therefore $\frac{E}{R_t} = \frac{E}{R_1} + \frac{E}{R_2} + \frac{E}{R_3}$

Divide through by E and you have,

$$\frac{1}{R_t} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

This equation is known as the reciprocal formula for finding the total or equivalent resistance of a parallel circuit. Another form of the equation may be obtained by solving for R_t. Thus,

$$R_t = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}$$

An analysis of the equations for total resistance in a parallel circuit reveals that R_t is always less than the smallest resistor in a parallel circuit. Thus, a 10-ohm, a 20-ohm, and a 40-ohm resistor connected in parallel have a total resistance of less than 10 ohms.

The greater the number of resistors in parallel, the less is the total resistance. Whenever a resistor is added in parallel, another path for current to flow is provided; this increases the total current and decreases the total resistance.

Two Resistors In Parallel

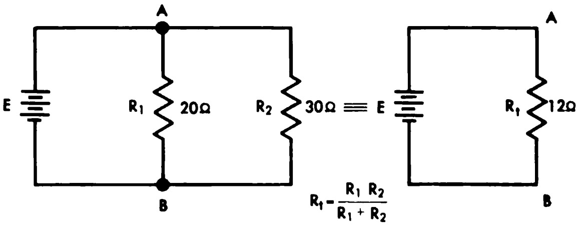
If there are only two resistors in a parallel circuit, the reciprocal formula is $\frac{1}{R_t} = \frac{1}{R_1} + \frac{1}{R_2}$. This may be simplified as follows:

$$\frac{1}{R_t} = \frac{R_2 + R_1}{R_1 \times R_2}$$

$$R_t = \frac{R_1 \times R_2}{R_1 + R_2}$$

The equivalent resistance of two resistors in parallel is equal to *the product of the two resistors divided by their sum*. For example, a 20-ohm resistor and a 30-ohm resistor in parallel are equivalent to a resistor of 12 ohms. This is found as follows:

$$R_t = \frac{20 \times 30}{20 + 30} = \frac{600}{50} = 12 \text{ ohms}$$



Equal Resistors In Parallel

The total resistance of any number of equal resistors connected in parallel is equal to *the resistance of one resistor divided by the number of resistors*. Expressed mathematically,

$$R_t = \frac{R}{N}$$

where R_t is the total resistance, R is the resistance of one resistor, and N is the number of resistors.

For three resistors, this equation may be derived as follows:

$$\frac{1}{R_t} = \frac{1}{R} + \frac{1}{R} + \frac{1}{R}$$

$$\frac{1}{R_t} = \frac{3}{R}$$

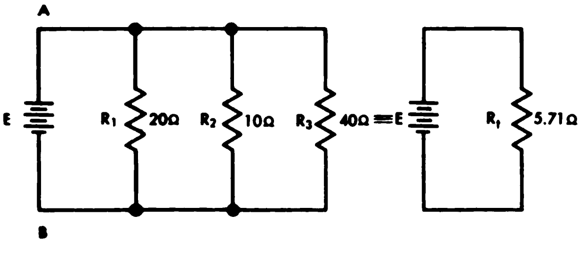
$$R_t = \frac{R}{3}$$

Two 10-ohm resistors connected in parallel are equivalent to 5 ohms. Five 30-ohm resistors connected in parallel are equivalent to a resistance of 6 ohms.

Assumed Voltage Method

To make it easier to find the equivalent resistance of any number of resistors in parallel, you may find it advisable at times to assume a convenient voltage across the parallel branch. The currents that would flow through each branch with the assumed voltage applied are then added to obtain the total current. The equivalent resistance is found by using Ohm's law — dividing the assumed voltage by the total current.

For example, you desire to find the equivalent resistance of the parallel circuit shown below. Assume a voltage that is easily divisible by the value of each of the resistors. In this case, assume 80 volts is impressed across points A and B. Therefore: $I_1 = \frac{80}{20} = 4$ amps, $I_2 = \frac{80}{10} = 8$ amps, and $I_3 = \frac{80}{40} = 2$ amps.



$$I_t = I_1 + I_2 + I_3$$

$$I_t = 4 + 8 + 2 = 14 \text{ amps}$$

$$R_t = \frac{E_t}{I_t} = \frac{80}{14} = 5.71 \text{ ohms}$$

POWER

The total power used in a parallel circuit is equal to the sum of the powers dissipated in each of the components of the circuit. This may be shown mathematically as follows:

$$I_t = I_1 + I_2 + I_3$$

$$E_t = E_1 = E_2 = E_3$$

Multiply both sides of the first equation by E (E_t, E_1, E_2, E_3).

$$E_t\ I_t = E_1\ I_1 + E_2\ I_2 + E_3\ I_3$$

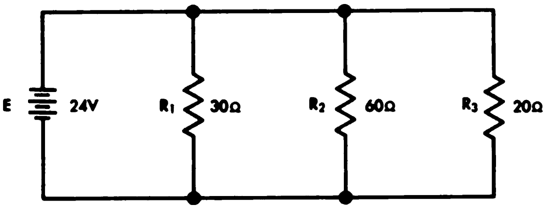
But $P = E \times I$

Therefore, $P_t = P_1 + P_2 + P_3$

ANALYZING PARALLEL CIRCUITS

The following examples will show how the equations derived for a parallel circuit may be used in analyzing circuit problems.

1. Three resistors are connected in parallel across a 24-volt battery. The values of the resistors are 30 ohms, 60 ohms, and 20 ohms. What is the equivalent total resistance? How much current flows through each resistor? How much is the total current?



Given: $E = 24$ volts, $R_1 = 30$ ohms

$R_2 = 60$ ohms, $R_3 = 20$ ohms

Find: R_t, I_1, I_2, I_3, I_t

Solution: (a) $\frac{1}{R_t} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$

$$\frac{1}{R_t} = \frac{1}{30} + \frac{1}{60} + \frac{1}{20}$$

$$\frac{1}{R_t} = \frac{2+1+3}{60} = \frac{6}{60} = \frac{1}{10}$$

$$R_t = 10 \text{ ohms}$$

(b) $I_1 = \frac{E}{R_1} = \frac{24}{30} = .8$ amp

$$I_2 = \frac{E}{R_2} = \frac{24}{60} = .4 \text{ amp}$$

$$I_3=\frac{E}{R_3}=\frac{24}{20}=1.2 \text{ amp}$$

(c) $I_t=\frac{E_t}{R_t}=\frac{24}{10}=2.4$ amp

or

$I_t=I_1+I_2+I_3$

$I_t=.8+.4+1.2=2.4$ amp.

2. In example 1, how much power is dissipated in each resistor? How much power is dissipated in the entire circuit?

Solution: (a) $P_1=E_1\times I_1$

$P_1=24\times .8$

$R_1=19.2$ watts in R_1

$P_2=E_2\times I_2$

$P_2=24\times .4$

$P_2=9.6$ watts in R_2

$P_3=E_3\times I_3$

$P_3=24\times 1.2$

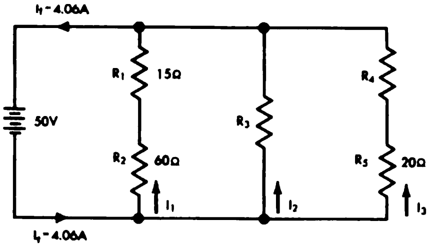
$P_3=28.8$ watts in P_3

(b) $P_t=E_t\ I_t$ or $P_t=P_1+P_2+P_3$

$P_t=24\times 2.4$ $P_t=19.2+9.6+28.8$

$P_t=57.6$ watts $P_t=57.6$ watts

3. In the parallel circuit below, find: (a) the power expended across R_1, (b) the current I_2, (c) resistance R_3, (d) the power expended across R_3, (e) resistance R_4, (f) the power expended across R_4 (g) total power expended in the circuit.



Given: $E_b=50$ volts, $I_t=4.06$ amps,
$I_3=1.6$ amp
$R_1=15$ ohms, $R_2=60$ ohms
$R_5=20$ ohms

Find: P_1, I_2, R_3, P_3, R_4, P_4, P_t

Solution: (a) $I_1=\frac{E_b}{R_1+R_2}$

$I_1=\frac{50}{75}=.66$ amps

$P_1=I_1^2R_1$

$P_1=(.66)^2\times 15=6.53$ watts

(b) $I_t=I_1+I_2+I_3$

$I_2=I_t-I_1-I_3$

$I_2=4.06-.66-1.6=1.8$ amp

(c) $R_3=\frac{E_b}{I_2}$

$R_3=\frac{50}{1.8}=27.7$ ohms

(d) $P_3=(I_2)^2\ R_3$

$P_3=(1.8)^2\times 27.7$

$P_3=89.7$ watts

(e) By Kirchhoff's Law

$E_b=I_3R_4+I_3R_5$

$I_3R_4=E_b-I_3R_5$

$1.6R_4=50-1.6\times 20$

$1.6R_4=50-32=18$

$R_4=\frac{18}{1.6}=11.25$ ohms

(f) $P_4=(1_3)^2R_4$

$P_4=(1.6)^2\times 11.25$

$P_4=2.56\times 11.25=28.8$ watts

(g) $P_t=I_tE_b$

$P_t=4.06\times 50$

$P_t=203$ watts

SUMMARY OF PARALLEL CIRCUITS

The voltage across each branch of a parallel circuit is equal to the voltage across every other branch and to the applied voltage.

The total current is equal to the sum of the currents in the parallel branches.

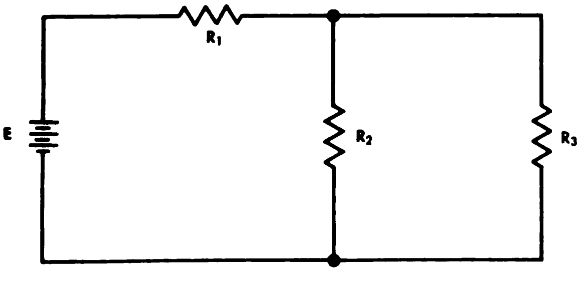
The reciprocal of the total (or equivalent) resistance of a parallel circuit is equal to the sum of the reciprocals of the resistances of the branches.

The total power dissipated is equal to the sum of the powers dissipated by the components.

Chapter 7

Series-parallel Circuits

Most circuits in electronic equipment are not series circuits or parallel circuits but are combinations of series and parallel circuits. A *series-parallel* circuit consists of groups of parallel resistors connected in series with other resistors. An example of a series-parallel circuit is shown below. Note that the total current flows through R_1 but divides through R_2 and R_3 in parallel.

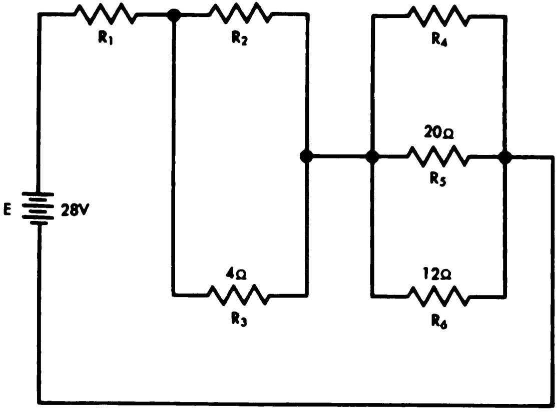


A Series-parallel Circuit

ANALYZING A SERIES-PARALLEL CIRCUIT

To find the total resistance and total current in the series-parallel circuit shown below, proceed as follows:

Reduce each group of parallel resistors to a single equivalent resistor and draw this equivalent circuit.



$$R_A = \frac{R_2R_3}{R_2 + R_3} = \frac{12 \times 4}{12 + 4} = \frac{48}{16} = 3 \text{ ohms}$$

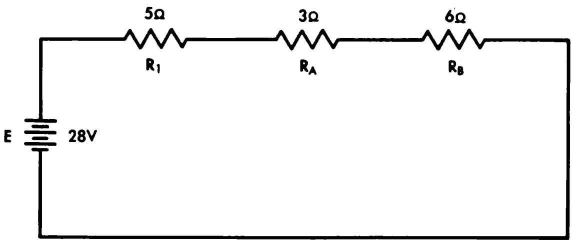
$$\frac{1}{R_B} = \frac{1}{R_4} + \frac{1}{R_5} + \frac{1}{R_6}$$

$$\frac{1}{R_B} = \frac{1}{30} + \frac{1}{20} + \frac{1}{12} = \frac{2 + 3 + 5}{60} = \frac{1}{6}$$

$$R_B = 6 \text{ ohms}$$

The equivalent circuit is shown below.

Use the equivalent circuit to solve for R_t and I_t.



Equivalent Series Circuit for Series-Parallel Circuit

$$R_t = R_1 + R_A + R_B$$

$$R_t = 5 + 3 + 6 = 14 \text{ ohms}$$

$$I_t = \frac{E}{R_t} = \frac{28}{14} = 2 \text{ amp}$$

To further analyze the series-parallel circuit to find the current through and the voltage across each resistor, proceed as follows:

Use the equivalent circuit and find the voltage across each resistor by Ohm's law,

$$E_1 = I_t R_1$$

$$E_1 = 2 \times 5 = 10 \text{ volts}$$

$$E_A = I_t R_A$$

$$E_A = 2 \times 3 = 6 \text{ volts}$$

$$E_B = I_t R_B$$

$$E_B = 2 \times 6 = 12 \text{ volts}$$

E_1 is the voltage across R_1, E_A is the voltage across R_2 and R_3 in parallel, and E_B is the voltage across R_4, R_5, and R_6 in parallel.

The current through each resistor is found by using Ohm's law with the original illustration

$$I_1 = I_t = 2 \text{ amps}$$

$$I_2 = \frac{E_2}{R_2} = \frac{6}{12} = .5 \text{ amp}$$

$$I_3 = \frac{E_3}{R_3} = \frac{6}{4} = 1.5 \text{ amp}$$

$$I_4 = \frac{E_4}{R_4} = \frac{12}{30} = .4 \text{ amp}$$

$$I_5 = \frac{E_5}{R_5} = \frac{12}{20} = .6 \text{ amp}$$

$$I_6 = \frac{E_6}{R_6} = \frac{12}{12} = 1 \text{ amp}$$

The power dissipated across each resistor can be determined from the information already obtained.

The procedure for solving a series-parallel circuit problem may be summarized as follows:

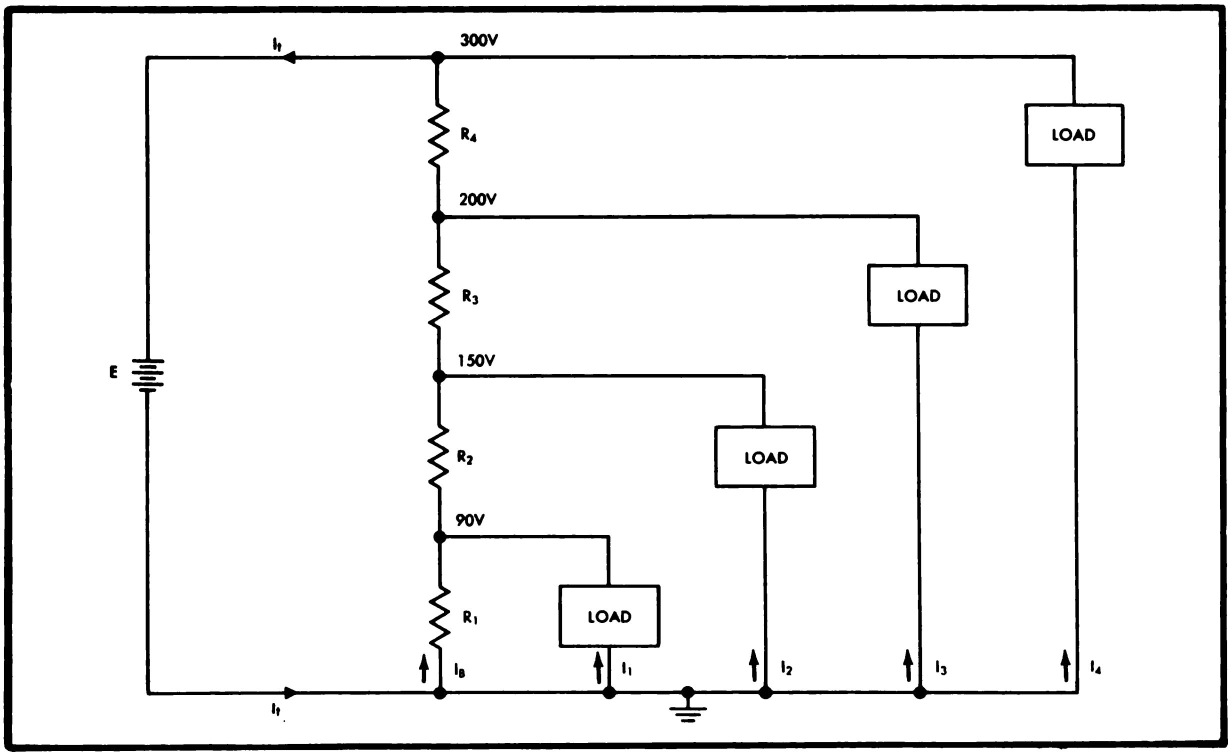
1. Reduce each parallel group to an equivalent resistance and redraw the circuit using these equivalent resistances.
2. Solve the resulting circuit for the total resistance and the total current.
3. Find the voltage drop across each parallel group.
4. Solve for any desired quantity.

VOLTAGE DIVIDERS

Unlike the situation in the home, where practically all electric appliances operate at the same voltage, radio and radar sets require several different values of voltage — yes, even the small five-tube AC/DC receiver requires more than one value of voltage. It is impractical to have a separate power source for each value of voltage especially since some electronic systems require dozens of different voltage values. A device which makes it possible to obtain more than one voltage from a single power source is known as a *voltage divider.*

A voltage divider usually consists of a resistor (or resistors connected in series) with fixed or movable contacts and two fixed terminal contacts. As current is made to flow through the resistor, different voltages may be obtained between the contacts. A typical voltage divider circuit is shown in the illustration on page 39.

A *load* is any device which draws current. A heavy load means a heavy current drain. In addition to the current drawn by the various loads, there is a certain amount drawn by the voltage divider itself. This is known as *bleeder current.* Bleeder current is desirable, but it is kept to a comparatively small value. It is common practice to choose resistor values for the voltage divider that will result in a *total* current flow of approximately 90% of the current rating of the power source. As an example: if a power source is rated at 100 ma., the voltage divider is



A Typical Voltage Divider

designed to cause a total current of approximately 90 ma. under load. If the load is 80 ma., the bleeder current must not exceed 10 ma.

To understand how a voltage divider works, examine the illustration above carefully and observe the following:

Each of the loads draws a given amount of current: I_1, I_2, I_3, I_4.

In addition to the above load currents, some bleeder current I_B flows.

The current I_t is drawn from the power source and is equal to the sum of *all* currents.

The voltage at each point is measured with respect to a *common* point as indicated by the ground symbol. Note that the common point is the point at which the total current (I_t) divides into separate currents (I_1, I_2, I_3, I_4, I_B).

Each part of the voltage divider has a different amount of current flowing in it. The current distribution is as follows:

Through R_1 — I_B
Through R_2 — $I_B + I_1$
Through R_3 — $I_B + I_1 + I_2$
Through R_4 — $I_B + I_1 + I_2 + I_3$

The voltage across each resistor of the voltage divider is:

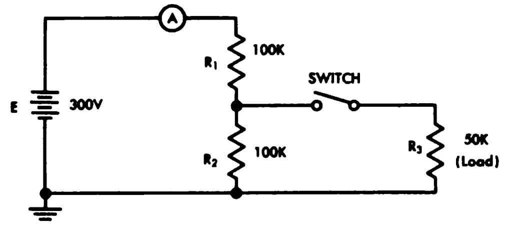
90 volts across R_1
60 volts across R_2
50 volts across R_3
100 volts across R_4

Since both the voltage and current are known for each resistor of the voltage divider, the resistance and power dissipated by each resistor may be easily calculated. In many applications, large amounts of current flow in the voltage divider. As a result, the resistors have a high wattage rating and are called *power resistors*.

The *power rating* of a resistor denotes the maximum amount of power the resistor can dissipate without being damaged. The power dissipated by each resistor may be used in selecting resistors of the proper power rating. However, if the voltage divider is a single tapped or adjustable resistor, the power rating is calculated using $P = I^2R$. The value of I used is the maximum current in any part of the resistor and the value of R is the total resistance of the voltage divider.

Effect of Changing Load

The effect of a change in load on the voltages obtained from a voltage divider may be illustrated by the following example. In the circuit shown below, the voltage obtained at point A with the switch open is equal to 150 volts. The current which flows through the circuit is equal to .0015 amp.



Effect of Changing Load

When the switch is closed, R_3 is connected in parallel with R_2. The parallel combination of R_2 and R_3 is equivalent to a resistance of 33.3K ohms. The voltage at point A is now equal to one-fourth the applied voltage or 75 volts. The total current is now equal to $\frac{300}{133.3K}$ or .002 amp.

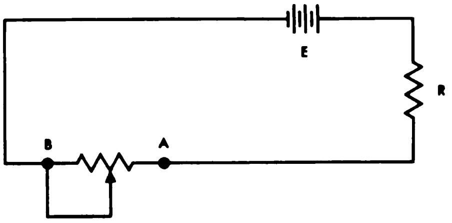
Changing the load in the circuit causes an increase in the total current and a corresponding decrease in the voltage at point A.

RHEOSTATS

Variable resistors, which have already been described on page 11, may be classified as rheostats or potentiometers, depending upon their use, the number of terminals, and how they are connected in the circuit.

A rheostat is a variable resistor which may be used as a control *to vary the amount of current* which flows in an electric device. The rheostat has two terminals; one terminal is connected to one end of the resistance and the other terminal, to the slider arm contact. A typical circuit in which a rheostat is used is shown in the illustration in the next column. The device shown could be an electric motor or the pilot light of a radio. Since the rheostat controls the circuit, it serves to vary the motor speed or the intensity of the pilot light.

Examine the schematic to see how the rheostat performs its function. As the slider arm is moved from A toward B, the amount of rheostat resistance (AB) in the circuit is increased. Since the rheostat resistance and the fixed resistance of the device are in series, the total resistance of the circuit also increases. The total current in the circuit therefore decreases. By a similar analysis, you can see that as the slider arm is moved toward A, the total resistance decreases and the current through the device increases.



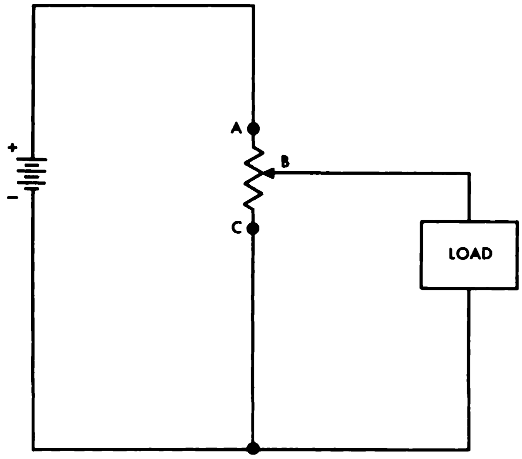
Controlling Current Flow with Rheostat

POTENTIOMETERS

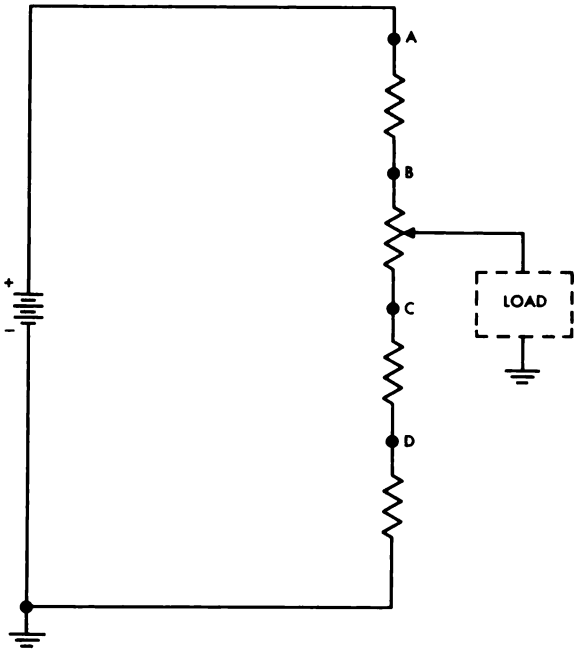
Like the rheostat, the potentiometer is also a variable resistor. However, the potentiometer has three terminals; the two ends and the slider arm are connected in the circuit. A potentiometer is a control used *to vary the amount of voltage* applied to an electrical device. By omitting the connection to one end of a potentiometer, you may use it as a rheostat. The potentiometer is one of the most common controls found in electronic applications. The volume control of your radio receiver and the BRIGHTNESS control of your television receiver are examples of its use.

In the circuit shown on page 41, note how the potentiometer is used as a means of obtaining a variable voltage from a fixed voltage source for an electrical device. The voltage applied across the device is equal to the voltage from B to C because the load is connected in parallel. When the slider arm is moved to point A, the entire voltage is across the load; when the slider arm is moved to point C, the voltage across the load is zero. The potentiometer makes it possible to apply any voltage within the range between zero and full voltage across the load.

Trace the current in the circuit. The current flows from the battery and divides into two parts. One part flows through the lower portion BC of the potentiometer, and the other part flows through the device. Both parts combine



Varying Voltage with a Potentiometer



A Potentiometer as Part of a Voltage Divider

and flow through the upper portion AB of the potentiometer back to the battery.

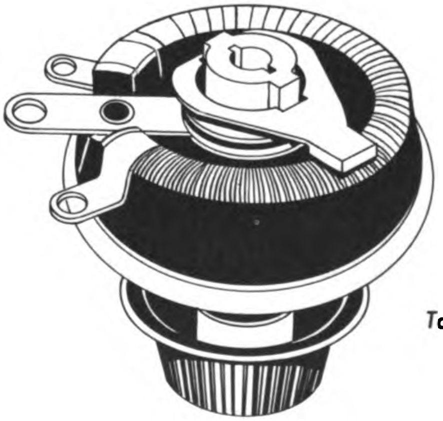
In choosing a potentiometer for a given purpose, be sure that it can withstand the current drawn by the load as well as the current it draws itself. The power rating of a potentiometer is calculated by using the formula $P = I^2R$, where I is the total current and R is the total resistance of the potentiometer.

Note that the useful work is done by the current through the load.. The energy in the current drawn by the potentiometer is dissipated in the form of heat. It is, therefore, advisable to keep this wasted current as small as possible by making the resistance of the potentiometer as large as is practicable. In most applications the resistance of the potentiometer should be several times the resistance of the load.

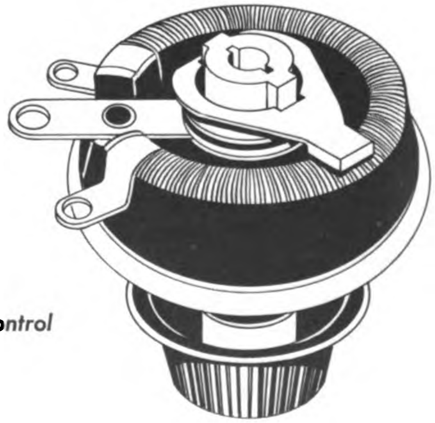
In addition to acting as a voltage divider to control the voltage across a load, potentiometers are often used as part of a voltage divider. The schematic shown to the right above shows a potentiometer used in this manner. The voltage across the load L can be varied by varying the potentiometer.

Linear and Tapered Controls

As described on page 11, rheostats and potentiometers are constructed of a resistance material over which a sliding contact moves. The resistance may be distributed in many ways;



Tapered Control



Linear Control

the type of distribution determines the classification of the control as *linear* or *tapered*. The linear control has its resistance evenly distributed over its entire length. The *tapered* control has more resistance per unit length at one end than it has at the other. As an example, a one-half turn of a linear control results in one-half of the total resistance between either end and the slider, while a one-half turn of a tapered control may result in one-tenth (or any fraction) of the total resistance between one end and the slider.

Many types of tapered controls are used in electronics, and each is chosen to perform a specific job. The volume control on your radio receiver is one type of tapered control. The taper is used to compensate for the fact that the human ear does not respond linearly to the intensity of sound. Normally, you expect a volume control to set the volume at half full volume when it is at the half-way mark, but actually, only one-tenth of the voltage is required to give the sensation of half full volume. Because of the above facts, most volume controls are constructed so that the midpoint of the potentiometer divides the resistance into one-tenth and nine-tenths of the total resistance. However, most potentiometers that are used to control voltages other than audio voltages are of the linear type.

Chapter 8

Magnetism

Some ancient Greeks discovered that a certain type of rock was capable of attracting iron and steel. This rock is now known as *magnetite*. The Chinese found that if a piece of magnetite were mounted in a horizontal plane and allowed to swing free, it would turn so that one end always pointed north. This fact was used later as an aid to navigation, and the stone became known as a leading stone or *lodestone*. An object which has the property of attracting iron and steel and which points to the north when allowed to swing free is called a *magnet*. Since lodestone has these properties when taken from the earth, it is called a *natural* magnet.

ARTIFICIAL MAGNETS

If a piece of iron or steel is rubbed or stroked with a lodestone, it acquires the properties of a magnet. A magnet made in this way is called an *artificial* magnet. However, this method of manufacturing magnets is impractical for commercial purposes because such magnets are very weak. Strong artificial magnets are now made commercially by passing a current through coils placed around the material to be magnetized. The principles which explain this phenomenon of electromagnetism will be discussed more fully later.

Artificial magnets may be either *permanent* or *temporary*. A permanent magnet is one which retains its magnetism for a relatively long time after it is made. Hardened steel will become a permanent magnet when magnetized. This property of a substance which causes it to remain magnetized is called *retentivity*. A temporary magnet is one which loses its magnetism almost immediately after the magnetizing force is removed. Soft iron is an example of a metal which is used as a temporary magnet. It has very little retentivity. Therefore, soft iron is an ideal core for electromagnets where it is desired that the magnetic characteristics be removed when current stops flowing.

MAGNETIC POLES

If a bar magnet is placed on a flat surface and sprinkled with iron filings, the iron filings will collect around it in the manner shown in the illustration below. Note that the iron filings cling to the magnet in clusters at the two ends and that the center has very few filings on it. The two ends, where the magnetic strength appears to be strongest, are called the *poles* of the magnet. The concentration at both ends appears to be the same, indicating that both poles are of equal strength.

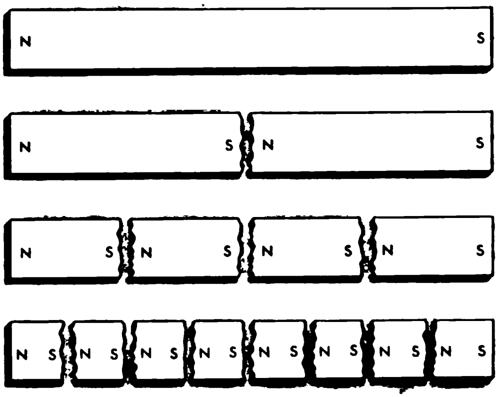
Poles of a Magnet

If a bar magnet or magnetized needle is suspended or pivoted, it comes to rest with its poles indicating a general north and south direction. The same pole always points to the north and is called the north-seeking pole or merely the *north* pole. The letter N on a magnet indicates this pole, and the opposite pole is labeled S, for south.

MOLECULAR THEORY OF MAGNETISM

If a bar magnet is broken in two, each part exhibits the properties of a magnet when dipped in iron filings or suspended to swing free. It appears that breaking the magnet creates two new poles at the point where the break occurs. Were this process of breaking the magnet continued a few more times, each fragment would behave like a magnet with two poles as shown in the next illustration. This action seems to indicate that the property of magnetism exists throughout the entire magnet although its strength seems to be concentrated at the poles.

Try to picture the magnet broken an infinite number of times into fragments so small that each fragment is a molecule. These molecules of magnetized steel can be considered to be very small magnets, each of which has two poles. The theory that a magnetic substance is composed of an infinite number of molecular magnets is called the *Molecular Theory of Magnetism*. According to this theory, the molecules of an unmagnetized bar of steel point in all directions as shown in A of the illustration at the top of page 45. This haphazard placement of molecules causes the north and south poles to neutralize each other. When the bar of steel is placed in a magnetic field, however, the molecules become rearranged in a complete chain of rows across the magnet with all north poles facing the same way. This is shown in B of the illustration.



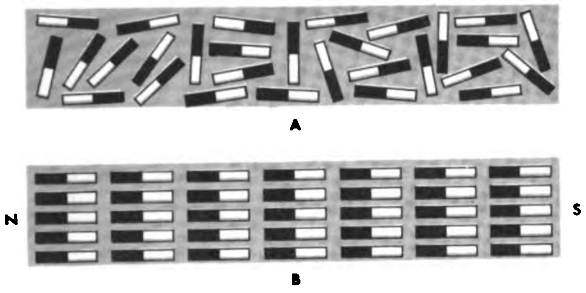
Effect of Breaking a Bar Magnet

This theory makes it possible to explain many observed facts in a logical manner. For example, hard steel requires a stronger magnetic field to produce a given degree of magnetization than soft iron because its molecular magnets are turned with more difficulty. However, when the magnetizing force is removed, the molecular magnets of the steel retain their position while those of the soft iron return to their original haphazard positions more readily.

The small amount of magnetism that remains in an electromagnet after the magnetizing force is removed is called *residual* magnetism. This is explained by the fact that all the molecular magnets have not yet returned to their original haphazard positions.

FORCE BETWEEN POLES

A bar magnet suspended in air and free to move will be attracted or repelled by another magnet brought close to it. If the poles that are brought close to each other are the same, the suspended magnet will be repelled; if they are different, the suspended magnet will be attracted.



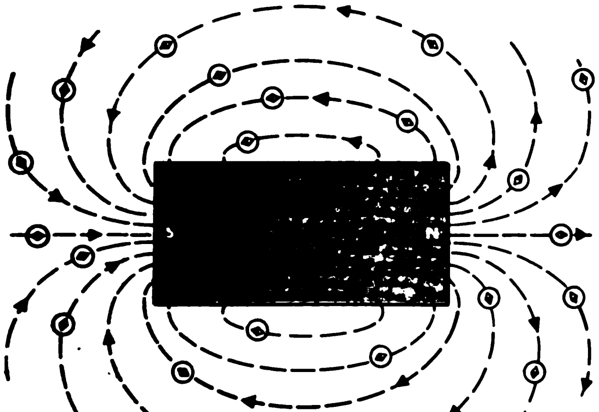
Bar of Steel A. Unmagnetized B. Magnetized

The magnetic laws of attraction and repulsion are stated as follows: *Like poles repel and unlike poles attract each other.*

The French physicist Charles A. Coulomb found that the force of attraction or repulsion between two magnets increased if the magnets were brought closer together and decreased if they were moved farther apart. He also found that this force increased as stronger magnets were used, and decreased as weaker magnets were used. The mathematical relationship between this force, the distance between the magnets, and the strength of the magnets may be stated as follows: *The force of attraction or repulsion between two magnetic poles varies directly as the product of the strength of the poles and inversely as the square of the distance between them.*

MAGNETIC FIELDS

The region around a magnet where its influence can be felt is called a *magnetic field.* The influence of a magnetic field can be observed by placing a piece of paper over a bar magnet and sprinkling iron filings over the paper. The pattern taken by these iron filings is shown in the illustration, Magnetic Field around a Bar Magnet shown below.



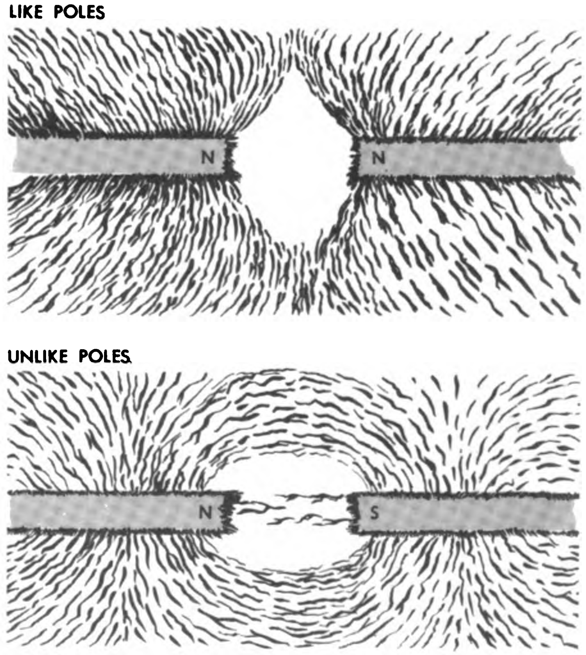
Magnetic Field around a Bar Magnet

If a second bar magnet is placed near this one and iron filings sprinkled over it, a pattern is obtained which depends on whether like or unlike poles are near each other. This is shown in the illustration below.

From the patterns that the iron filings take in the magnetic field, you can observe that the magnetic field takes the form of lines arranged in an orderly fashion. Although these lines are actually invisible, their effect on the iron filings or on a compass needle placed in the vicinity can be observed. These invisible lines are called *lines of magnetism* or *lines of induction.* One line is called a *maxwell.* The total number of lines of induction entering or leaving a pole is called its magnetic *flux.* The Greek letter ϕ (phi) is the symbol used for magnetic flux.

Magnetic lines have the following characteristics:

1. They always form complete loops. They travel from north to south outside the magnet and from south to north inside.
2. They never cross each other.


LIKE POLES

UNLIKE POLES

Magnetic Field between Two Bar Magnets

3. They expand or contract like rubber bands as a force is exerted upon them and return to their original state when the force is removed.

4. They will pass more easily through a magnetic material than through air or a nonmagnetic material. No material can completely prevent their passage.

FIELD INTENSITY

From the patterns the iron filings make and from the position a compass needle assumes as it is moved about within a magnetic field, it is evident that a magnetic field has a certain direction and particular intensity at every point.

The *direction* of a magnetic field is the direction of the force it exerts upon an isolated north pole. The *strength* of a magnetic field (field intensity) at any point is the force that it exerts upon a unit north pole placed there. The symbol used for field intensity is H and the unit is called the *oersted.* An oersted is the intensity of a magnetic field in which a unit pole experiences a force of one dyne.

Field intensity is represented by lines and denotes force per unit pole. These lines are therefore called *lines of force.* Lines of force and lines of induction are equal when the medium is air. For this reason, the term lines of force is often used for lines of induction. When the medium is a magnetic material, however, the lines of force and the lines of induction are not equal.

FLUX DENSITY

Magnetic flux was defined as the total number of lines of induction (maxwells) leaving or entering a magnetic pole. The total flux, however, does not give a true picture of the concentration of these lines. The area covered must also be considered. The number of lines of induction per square centimeter in a plane perpendicular to the direction of the magnetic field is called *flux density.* The symbol used to designate flux density is B and the unit is the *gauss.* A gauss is equal to 1 maxwell per sq. cm.

The lines of force (H) and the lines of induction (B) are equal in a vacuum (or air) but differ in a magnetic substance with B much larger than H. The ratio B:H can therefore be used to differentiate between magnetic and nonmagnetic substances. This ratio may be used to express — the ease with which magnetic lines can pass through a substance — *permeability.* Expressed as a formula

$$\mu = \frac{B}{H}$$

where μ (the Greek letter mu) is the permeability, B is flux density in gausses, and H is field intensity in oersteds.

CLASSIFYING MAGNETIC SUBSTANCES

Substances have been considered as either magnetic or nonmagnetic in the discussion to this point. However, if you consider a substance as magnetic if it is acted upon by a magnetic field, then all substances are magnetic to some degree. The permeability of a substance indicates the degree to which the substance is magnetic.

All substances react to a magnetic field in one of two ways. They are therefore classified as *diamagnetic* or *paramagnetic.* When made into a rod, a diamagnetic substance is repelled by a magnetic field and will align itself at right angles to the field. On the other hand, a paramagnetic substance, when made into a rod, will be attracted by a magnetic field and will align itself parallel to the field.

Diamagnetic substances have permeabilities slightly less than unity. Bismuth has the lowest permeability known (0.9998). Other diamagnetic substances are phosphorus, antimony, flint glass, and mercury. Such substances are usually considered to be nonmagnetic.

Paramagnetic substances have permeabilities greater than unity. Those whose permeabilities are only slightly greater than unity, like platinum (1.00002), are merely called *paramagnetic.* Such substances are usually considered to be nonmagnetic. Paramagnetic substances whose permeabilities are much greater than unity are called *ferromagnetic.* These are usually considered to be magnetic. Although the term ferromagnetic originally referred only to iron, nickel, and cobalt, many alloys (some of which may contain none of these elements) are now classified as ferromagnetic. Some such alloys are: permalloy, perminvar, alnico, and the Heusler alloys.

Permalloy yields a large flux density (B) for a very small field intensity (H). Perminvar is an alloy of iron, nickel, and cobalt. Alnico is made of aluminum, nickel, and cobalt alloyed

DIAMAGNETIC		PARAMAGNETIC		FERROMAGNETIC	
Antimony	0.999952	Aluminum	1.000022	Cobalt	170
Bismuth	0.999824	Erbium	1.00289	Nickel	1,000
Copper	0.99999	Oxygen (Gas)	1.0000019	Permalloy	1,000,000
Nitrogen (Gas)	0.999999995	Air	1.00000029	Perminvar	2,000
Water	0.9999992			Steel	1,500
				Iron	7,000

with iron. It has high retentivity which is essential for permanent magnets. The Heusler alloys are alloys of copper, manganese, and aluminum, none of which are ferromagnetic in themselves.

The permeabilities of some diamagnetic and paramagnetic substances as well as the maximum permeabilities of some ferromagnetic substances are shown in the table above.

MAGNETIC SHIELDING

Magnetic shielding is used to keep stray magnetic fields away from circuits and components that may be affected by these magnetic fields. Either of two methods may be used. In one method, the shield is made of a magnetic material such as soft iron or nickel. A container made of this material is placed around the circuit or component to be protected. This container then provides an easy path for stray magnetic lines of force around the protected circuit. The shield also keeps magnetic lines of force produced within the container from affecting circuits around it for the same reason.

The second method of shielding employs a shield made essentially of a nonmagnetic conducting material such as copper, tin, or aluminum. The stray magnetic fields are prevented from affecting the shielded circuit because of the manner in which they affect the shield. The stray magnetic field causes currents to flow in the conducting shield which set up magnetic fields that oppose or cancel the stray fields. How currents set up magnetic fields and how these magnetic fields cause currents to flow are explained in the discussion of electromagnetism which follows.

ELECTROMAGNETISM

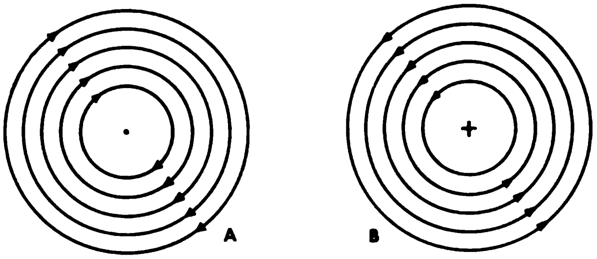
Hans C. Oersted, a Danish physicist, showed that there was a relationship between electricity and magnetism. He found that a conductor carrying current was surrounded by a magnetic field. He also discovered that the direction of the magnetic field at any point in the field was always tangent to a circle drawn around the conductor through the point. By measuring the strength of the magnetic field at various distances from the conductor, Oersted found the field was stronger closer to the conductor — and weaker farther from the conductor. Upon varying the current through the conductor, Oersted discovered that the strength of the magnetic field increased with an increase in current and decreased with a decrease in current.

Magnetic Field around a Wire

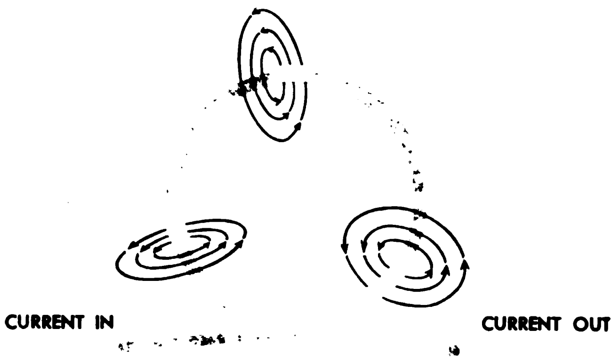
You can discover for yourself what Oersted did. Place a compass needle near a straight section of wire carrying current. The needle will indicate the presence of a magnetic field. As you move the compass around the wire, you will find that the magnetic field extends all around the wire and that the compass needle always points in the same direction (clockwise or counterclockwise) around the wire. If you stop the current flow, the needle returns to its normal position. If you reverse the direction of current flow and repeat the process, the compass needle will point in the opposite direction as you move it around the wire.

The lines of force around a conductor carrying current (as determined by the direction in which the compass needle points) travel in a clockwise or a counterclockwise direction depending upon the direction of current flow. At A in the first illustration on page 48, the electrons are moving out of the paper (as represented by the dot in the center). Note that the lines of force are shown in a clockwise direction. At B in the illustration, the electrons are moving into the paper (as represented by the +) and the lines of force are shown in a counterclockwise direction.

To determine the direction of lines of force around a conductor carrying current, use the



Directions of Lines of Force around a Wire Carrying Current



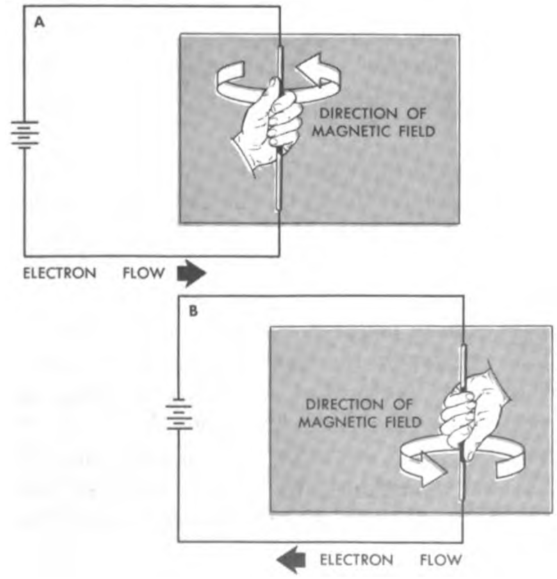
Flux in a Circular Loop

left-hand rule as follows: If you grasp a conductor with your left hand in such a manner that your thumb points in the direction of electron flow, your fingers will indicate the direction of the lines of force as shown below.

Vary the distance of a compass needle from the wire. Note that as you move it closer, the needle is held in position more firmly, and as you move it farther away, the magnetic field seems to have less effect. The compass needle is affected in the same way if you first increase and then decrease the current through the conductor.

Magnetic Field around a Solenoid

If you bend the straight conductor into a single turn loop, the lines of force will concentrate within the loop. This effect is shown in the illustration, Flux in a Circular Loop. The concentration within the loop is due to the fact that all the lines of force enter the loop from one side. If the direction of current flow were changed, all the lines of force would enter the loop from the other side, but the concentration would still be in the center.



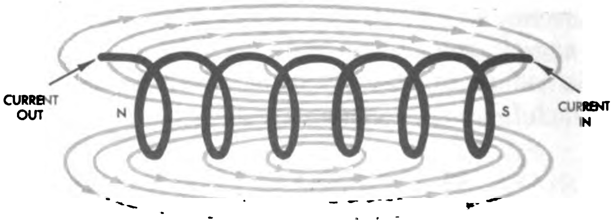
Left-hand Rule for Direction of Lines of Force

When a straight wire is wound into the form of a coil and current passed through it, the magnetic fields about all turns will have the same direction. The lines of force will be concentrated in the center of each turn.

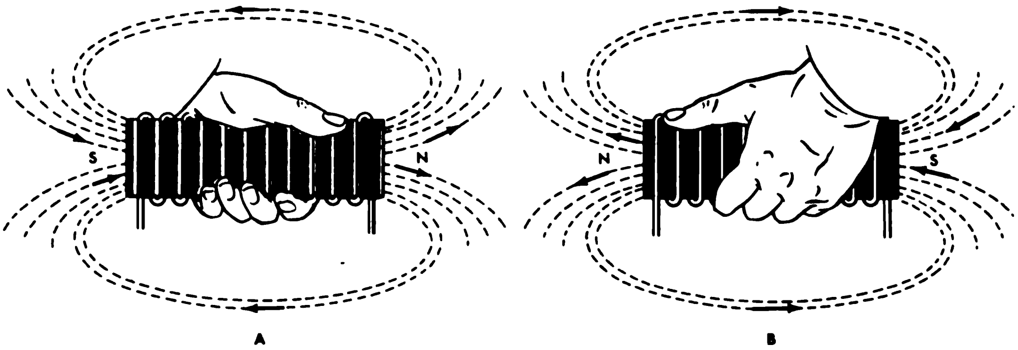
The magnetic field about a coil or solenoid, through whose windings a current is flowing, is shown below. Note that the magnetic field around the solenoid is similar to that around a bar magnet. To determine the polarity of a solenoid, use the *left-hand* rule for solenoids as follows: If you grasp the solenoid with the left hand so that your fingers point in the direction of electron flow, your thumb points toward the north pole, as shown on page 49.

ELECTROMAGNETS

An electromagnet is an application of a solenoid. To increase the strength of the magnetic field set up by an electric current, a soft iron core is inserted in the coil of the electromagnet. The additional magnetic strength is caused by the high permeability of soft iron; that is, magnetic



Magnetic Field around a Solenoid



Left-hand Rule for Solenoids

lines can pass readily through iron. By inserting a soft iron core into a solenoid, you greatly increase the number of magnetic lines of force. In addition to high permeability, soft iron has low retentivity; in other words, it retains its magnetism only while current is flowing.

Electromagnets are built in various sizes and shapes. Electromagnets of tremendous strength and size are used for unloading railroad cars, and small electromagnets are used in the bell, the telephone, etc. Electromagnets are of importance in much electrical equipment. Relays, meters, motors, generators, and loudspeakers are a few examples of electrical equipment in which electromagnets play an important part.

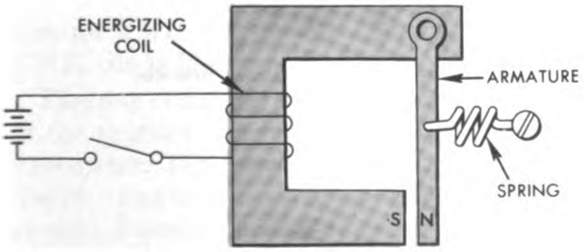
RELAYS

A relay is an electromagnet used to control electrical devices in other circuits or in the same circuit. A change in current flowing in the relay circuit produces an armature movement that opens or closes contacts which produce a change in electrical conditions. The basic components of a relay are shown in the illustration.

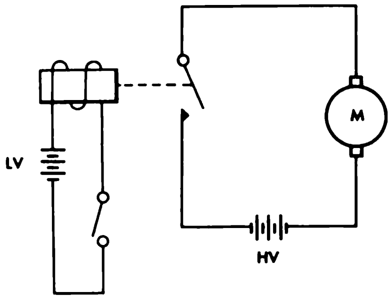
The magnetic circuit is composed of an electromagnet with part of its core pivoted or hinged. This pivoted portion is called the *armature* of the relay. An *energizing coil* through which current flows to produce the magnetic field and a *spring* are essential components of the relay. When the switch is closed, the electromagnet becomes energized and the armature is attracted toward its opposite magnetic pole. When the switch is opened, the electromagnet is de-energized and the spring pulls the armature back to its original position. Contact points are added to the armature and to the core to make better connections possible. Some relays have armatures with contact points above and below so that when one circuit is broken another is closed.

Starting Relay

To start a piece of equipment located at some distance, a starting relay may be used. In the circuit shown below, the closing of the switch makes an electromagnet out of the coil which then attracts the soft iron armature. The contact which is made completes the circuit to the motor. When the switch is opened, the mag-



Relay



Starting Relay Circuit

netic field collapses and the spring pulls the armature back to its original position breaking the circuit to the motor. In this way a relay and a switch in a circuit which draws very little current may be used to start a circuit in which there is high current drain. Long leads may be used in the relay circuit making it possible to keep the leads in the motor circuit short. This practice cuts down the expense involved.

A similar relay may be used to start circuits in which high and dangerous voltages are involved. Here the relay is used as a safety device. The switch closes a circuit in which only low voltages are used. The relay closes the high voltage circuit in which arcing may occur. Bodily harm to the person operating the switch is thus avoided.

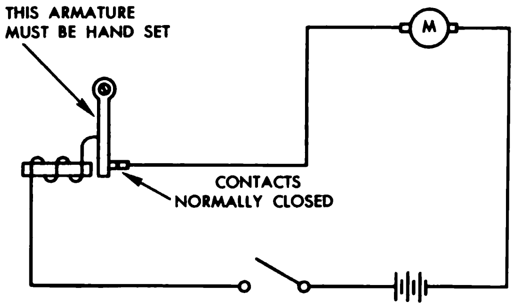
Overload Relay

An overload relay is a relay connected in series with a line. If the amount of current flowing in the line becomes excessive, the relay becomes energized and breaks the circuit. A circuit in which an overload relay is in series with a load is shown below. When the current becomes excessive, the pull of the magnet overcomes the tension of the spring and draws the armature toward the core. This causes the contacts to open and break the circuit. Once the circuit has been broken, the contacts must be reset manually in most overload relays.

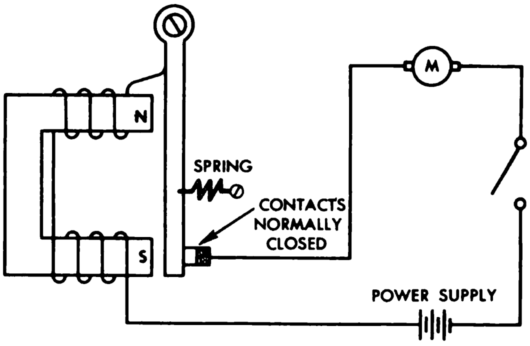
Circuit Breakers

A circuit breaker is like an overload relay or a fuse. It protects a circuit against short circuits and overloading. One type of circuit breaker consists of an electromagnet whose windings are in series with the load to be protected and with switch contact points. Excessive current causes the magnet to become energized and trip the switch which breaks the circuit to the circuit breaker and load. When the circuit fault has been corrected, the circuit breaker is reset.

An *interrupter* type circuit breaker is one in which the circuit is momentarily broken and then closed again. This action continues until the trouble is cleared. In the illustration below, the electromagnet becomes strong enough to attract its armature when the current becomes excessive. This breaks the circuit and the electromagnet becomes de-energized. A spring immediately closes the circuit again, and this action of making and breaking the circuit continues until the trouble is cleared. The alternate making and breaking of the circuit results in a buzzing sound which may serve as an alarm. Note that if the spring were removed and a switch provided for hand setting the contacts, this interrupter type circuit breaker would be a manual circuit breaker.



Overload Relay Circuit



Interruptor-type Circuit Breaker

Chapter 9

Electromagnetic Induction and Generators

In 1831, Michael Faraday made a discovery which proved to be of vast importance in the generation of electricity. In one of his experiments, he wound two different coils on an iron ring as shown on page 52. He connected a battery and key in series with one coil, and a galvanometer (a device detecting small electric currents) in series with the other. When he closed the primary circuit (P), the galvanometer in the seconday circuit (S) deflected momentarily and returned to zero. When he broke the primary circuit, the galvanometer again deflected momentarily (in the opposite direction) and returned to zero. Each momentary deflection indicated that current flowed in the secondary circuit and that therefore an EMF was present. Faraday called this EMF, the *induced EMF*.

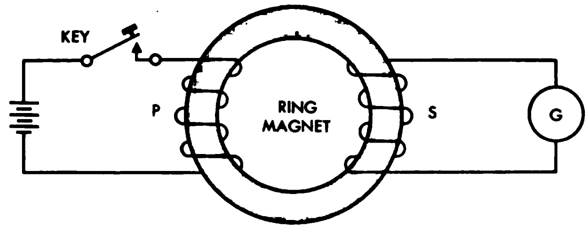
Faraday observed that a steady flow of current in the primary circuit had no effect on the galvanometer. The galvanometer deflected only at the moment of closing or opening the primary circuit. Faraday reasoned, therefore, that the induced EMF was caused by the *change* in the number of magnetic lines of force which linked the two coils.

INDUCING AN EMF

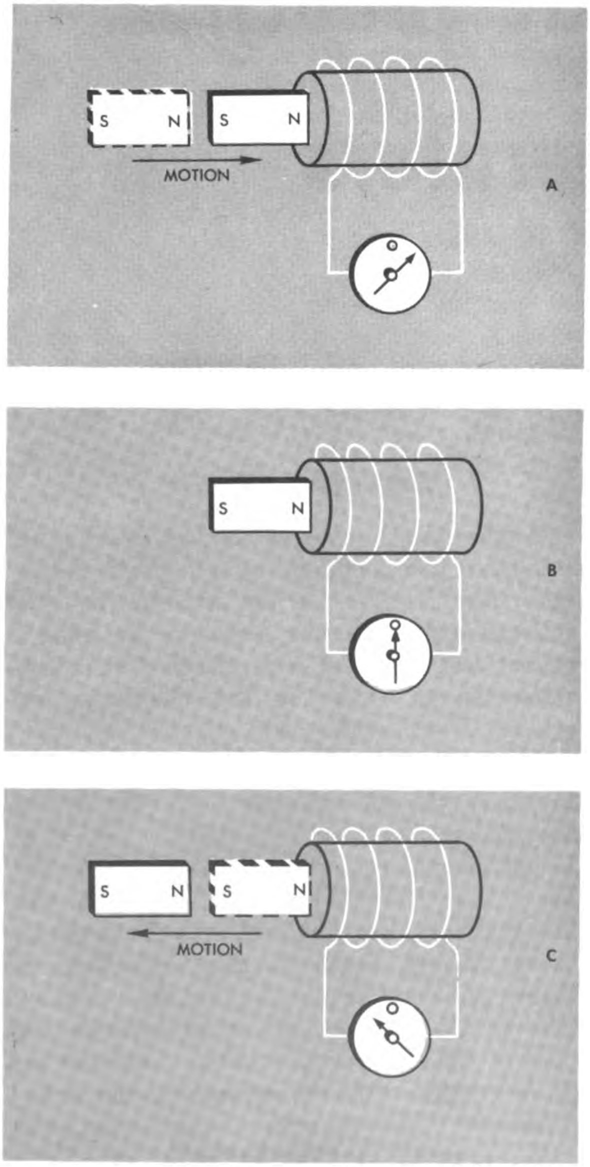
Faraday also discovered that an EMF could be induced in a conductor by a magnetic field if there were relative motion between the two. You can make a similar discovery by using a magnet and a coil of wire attached to a galvanometer as shown in the illustration on page 52.

As long as both the magnet and coil are kept stationary, the galvanometer reads zero. If you move the magnet toward the coil as shown at A, the galvanometer needle is deflected—an indication that an EMF is being induced in the coil. Stop the movement (diagram B) and the needle returns to zero. Move the magnet back to its original position (diagram C), and the needle deflects in the opposite direction. Increase the speed of the movement, and the deflection becomes greater. Stop the movement and it becomes zero.

The same effects can be produced if you move the coil and keep the magnet stationary. The deflection is greatest when the coil is moved *perpendicular* to the magnetic field and is zero when the coil is moved *parallel* to it. The speed



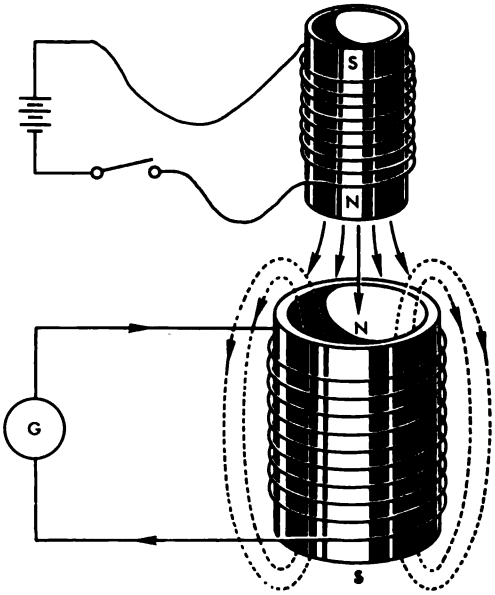
Faraday's Experiment



Inducing an EMF
with a Coil and a Magnet

of movement will affect the amount of deflection as before.

Replace the magnet by a solenoid in which current is flowing, as shown in the illustration below. Movement of the solenoid with respect to the coil or of the coil with respect to the solenoid will cause deflection of the galvanometer just as the magnet did.



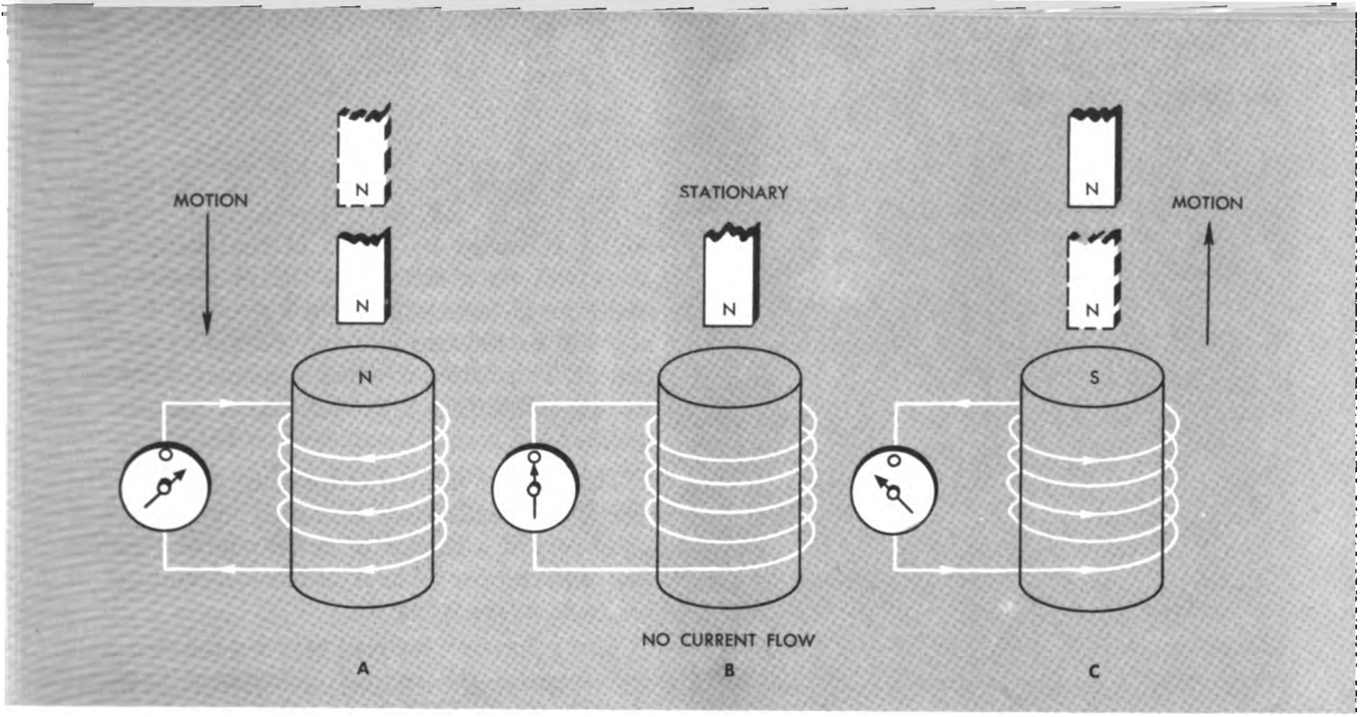
Inducing an EMF with a Coil and a Solenoid

From the above examples you may conclude, as Faraday did, that an EMF is induced in a coil of wire whenever there is a *change* in the magnetic flux linked with it. The magnitude of the induced EMF is determined by the *rate of change of magnetic flux linkages*. This conclusion is sometimes stated as follows: An EMF will be induced in a coil of wire if it cuts lines of force; the greater the number of lines of force the coil cuts per second, the greater is the induced EMF.

To increase the EMF induced in a coil by a magnetic field, it is necessary to make more flux linkages — cut more lines of force — per second. This result can be accomplished in each of these three ways:

1. Increasing the magnetic strength thus increasing the number of lines of force.

2. Increasing the speed of relative motion between the coil and the magnet.



Lenz's Law

3. Increasing the number of conductors by making the coil out of more turns of wire.

POLARITY OF INDUCED EMF

After performing experiments on induced EMF, the Russian physicist, H. F. Emil Lenz, arrived at a generalization called Lenz's law. This law states that, whenever a voltage is induced, polarity, is always such as to set up a current whose magnetic field will *oppose* the change of flux which induces the voltage.

A better understanding of Lenz's law may be obtained by studying the illustration, Lenz's Law shown above. As the north pole of the magnet is moving toward the coil, a voltage is induced in the coil which causes current to flow in such a direction as to create a *north* pole at the top of the coil. Thus the magnetic field created opposes the approach of the magnet. When the magnet's motion is stopped, the galvanometer indicates no current flow and no magnetic field exists around the coil. As the magnet is drawn upward, the voltage induced will be in such a direction as to cause the current to create a *south* pole at the top of the coil. Thus the magnetic field created opposes the pulling away of the magnet.

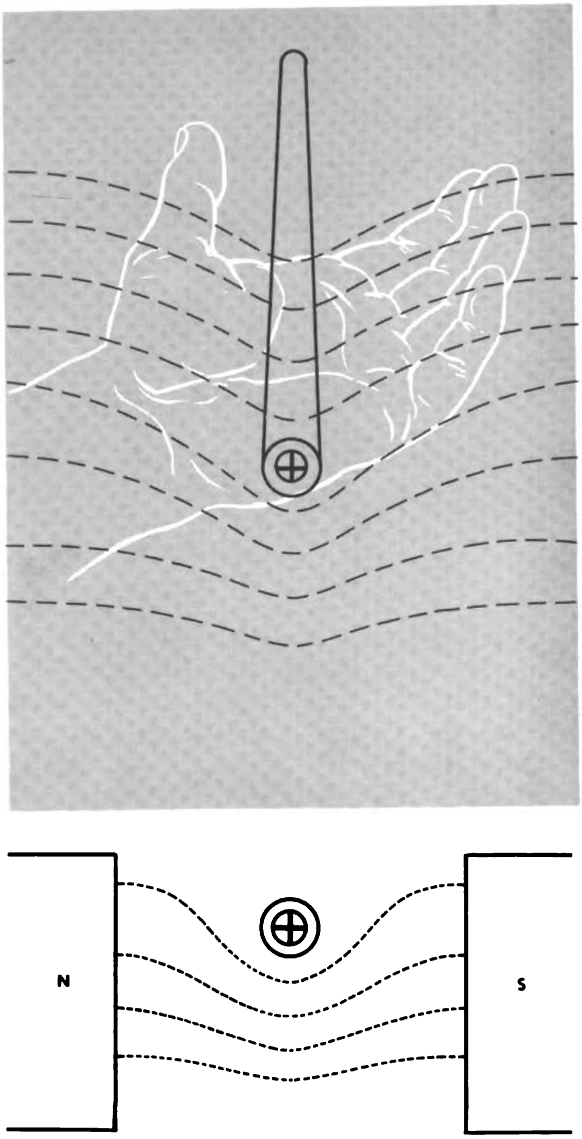
Lenz's law may be used to determine the direction of the electron flow induced in a wire which is moving perpendicular to a magnetic field as shown in the illustration on the next page. The electron flow in the wire is in such direction as to oppose the motion of the wire. This means that the magnetic field created must reinforce the original magnetic field at the bottom and oppose it at the top. A current flow into the page will create such a magnetic field causing a distortion of the magnetic field as shown. A simple rule for determining the direction of electron flow through the wire is: If you grasp the wire with the left hand so that the fingers follow the distorted lines caused by the motion of the wire, the thumb will point in the direction of electron flow.

Note

> The three directions — the *motion* of the wire, the *flux* lines, and the *electron* flow — are mutually perpendicular. If any two directions are known, the third may be determined.
>
> The motion of the conductor means the relative motion of the conductor with respect to the magnetic field. Either or both may be moving.

MUTUAL INDUCTION

A change of current in a conductor is always accompanied by a change in the magnetic field surrounding the conductor. An increase in current causes the field to *expand;* that is, the field intensity at any point increases. A de-



Current Flow in Wire Moving through a Magnetic Field

crease in current causes the field to *collapse* or decrease in intensity. If you place a conductor within an expanding or collapsing magnetic field, lines of force are cut and a voltage is induced in the conductor.

Suppose two coils are placed side by side as shown in the illustration, Mutual Induction. A battery and a rheostat are in series with coil A, and a galvanometer is connected to coil B. As long as the current in coil A is steady, there is no changing magnetic field and no voltage is induced in coil B. If the current through coil A is varied by means of the rheostat R, there will be a change in the magnetic field, and lines of force will be cut by coil B. A voltage will be induced and the galvanometer will deflect. This action in which a change of current in one coil causes a voltage to be induced in another coil is called *mutual induction.*

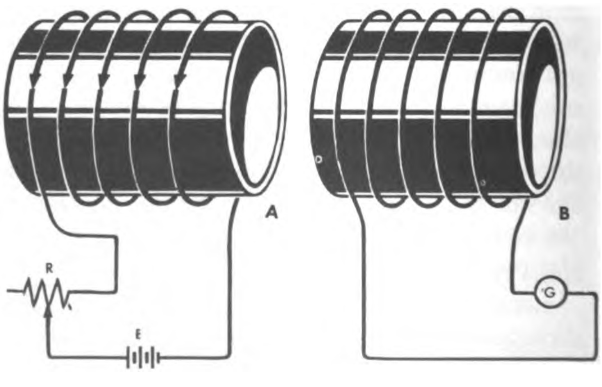
The unit of mutual inductance is called the *henry.* Two coils have a mutual inductance of one henry if a change of current in one coil of one ampere per second induces a voltage of one volt in the other.

A soft iron core inserted between the two coils will increase the mutual inductance between them because of the increase in the number of lines of force which pass through the soft iron.

An increase in the current through coil A causes an expanding field which is equivalent to moving the coil toward coil B. The voltage induced will (according to Lenz's law) cause a current to flow in coil B which is opposite in direction to the current in coil A. Similarly, a decrease in the current through coil A causes a collapsing field which induces a voltage and a current to flow in coil B in the same direction as in coil A.

SELF INDUCTION

As the current was changed in coil A, a change occurred in the flux and a voltage was induced in coil B. However, the changing flux also induced a voltage in coil A. This action in which a change in current induces a voltage in the same coil, because of the changing magnetic field, is called *self induction.*



Mutual Induction

This induced voltage is of such polarity as to oppose the change of current which caused it, and is called the *counter EMF* or *back EMF*.

The unit of self inductance is the henry. It is defined as the inductance of a coil in which a change of current of one ampere induces a counter EMF of one volt. Increasing the number of turns or inserting an iron core increases the inductance of a coil.

When two coils are close together, the relation between the mutual inductance and the individual inductances is equal to,

$$M = k\sqrt{L_1 \cdot L_2}$$

where k is the coefficient of coupling and has a value between 0 and 1. When all the flux produced by the current through one coil links all the turns of the other, $k = 1$.

FLUX LINKAGES

Flux linkages were treated in the discussion of Faraday's experiments as the number of lines of force being cut. Below is a further explanation of flux linkages.

If a current flows in a conductor, a magnetic flux is set up about the conductor. The flux encircles the conductor, and the current encircles the flux as shown in the illustration below. Note that the current and the flux it produces completely encircle and link each other. This is similar to the way in which the keys on a key chain link the chain, and the chain, in turn, links the keys.

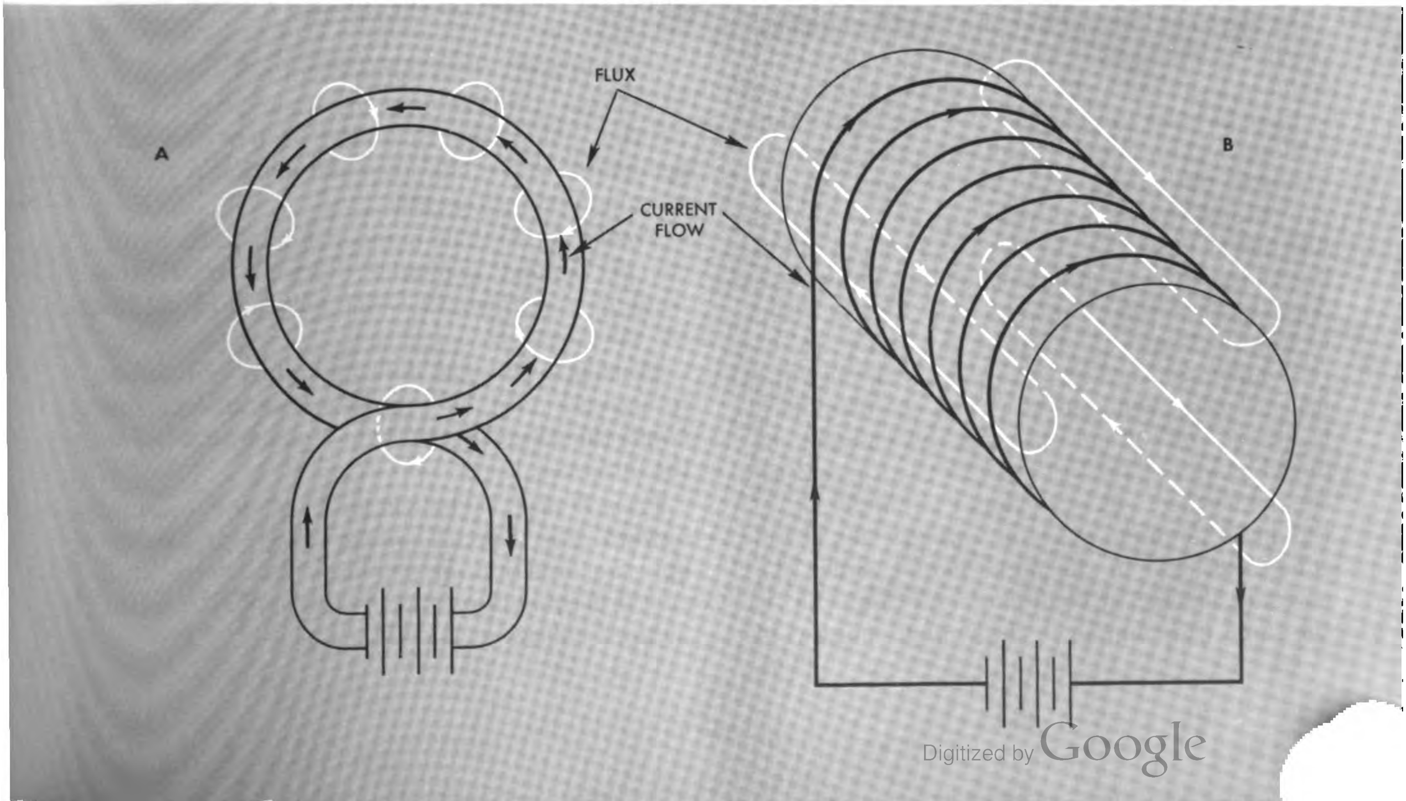
The product of the number of turns of a conductor and the number of lines of flux linking these turns is called the *flux linkages* of the circuit. The number of flux linkages per unit current is called the inductance of the circuit.

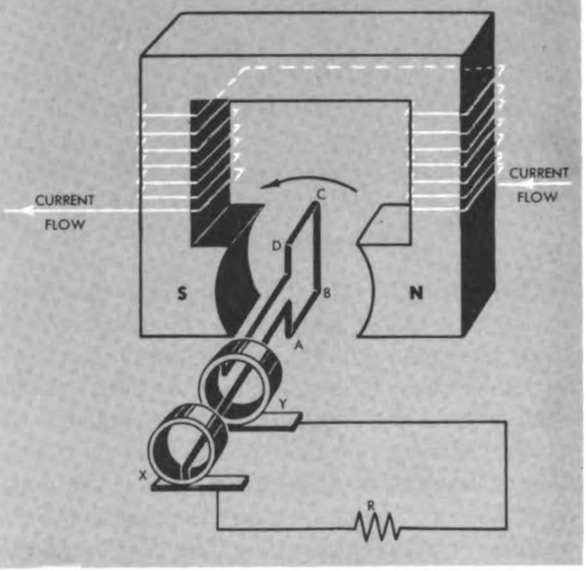
GENERATORS

The principles of electromagnetic induction are applied in the generation of electricity. As you have seen, a voltage is induced in a conductor if there is relative motion between the conductor and a magnetic field. A machine in which the magnetic field is provided by a permanent magnet is called a *magneto;* if the magnetic field is provided by an electromagnet, the machine is called a *dynamo*. The portion of the machine which contains the conductors in which the induced currents flow is called the *armature*. The motion may be obtained by rotating the armature, the magnetic field, or both.

A *generator* is a magneto or dynamo which changes mechanical energy to electrical energy. A *motor* is a magneto or dynamo which reverses the process and changes electrical energy to mechanical energy or motion. In many cases the same machine may be used as either a motor or a generator.

Flux Linkage





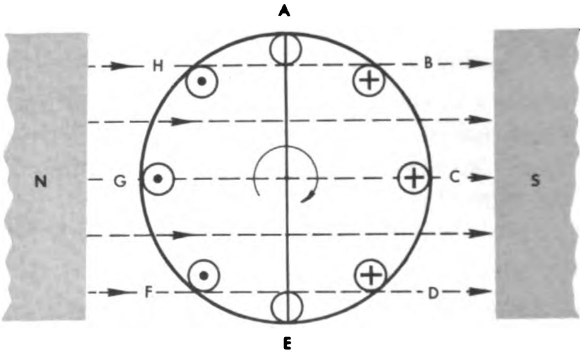
A Simplified AC Generator

A Simple AC Generator

A simplified generator is shown in the illustration above. It consists of a rectangular coil (armature) rotating between the poles of an electromagnet. AB and CD may be considered to be two conductors at the bottom and top of the coil connected to two rings called *slip rings* or *collector rings*. These rings are mounted with and rotate with the armature. They make contact with the outer circuit through brushes X and Y.

As the coil rotates about an axis perpendicular to the magnetic field, it cuts lines of force and induces a voltage. In the position shown, no voltage is being induced because the conductors AB and CD are moving parallel to the magnetic field and no lines of force are being cut. As the coil rotates counterclockwise, AB moves up and CD moves down. By the left-hand rule, you can determine that the induced voltage causes electrons to flow from A toward B in AB and from C toward D in CD. Through the outer circuit (R), current flows from Y to X back through the brushes to the conductors. Thus a complete path is formed for the current with the voltages induced in AB and CD in series. As the coil reaches the point where AB is moving down and CD up, the direction of the electron flow reverses in the conductors and in the outer circuit. Since the currents reverse in conductors AB and CD at the same time and since the voltage induced in AB is in series with the voltage induced in CD, you may consider them as one conductor for purposes of explanation.

Study an end view of this conductor taken at 45° intervals as the coil is rotated. The illustration below represents the conductor rotating in a clockwise direction in a magnetic field whose flux lines are from left to right.



End View of Conductor Rotating CW in a Magnetic Field

At A. The conductor is traveling parallel to the flux lines; no lines are being cut and the induced voltage is zero.

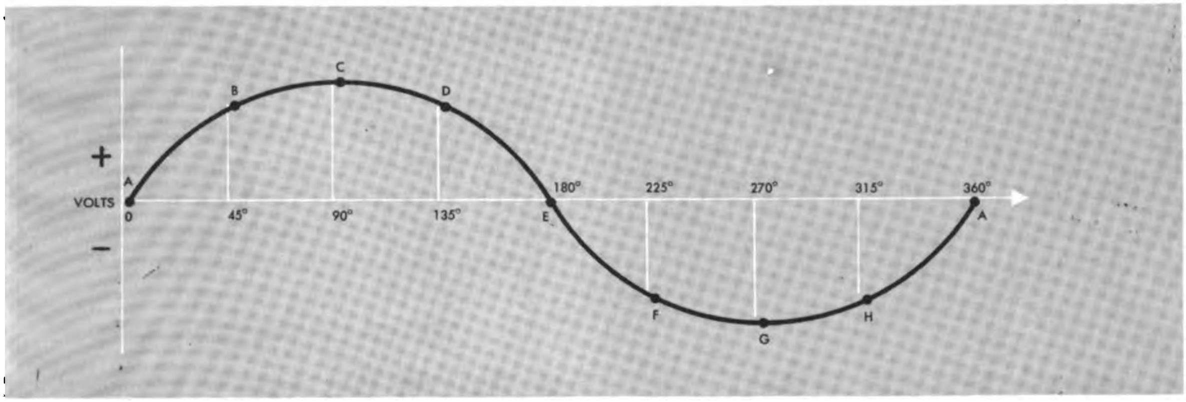
From A to B. As the conductor travels from A to B the number of flux lines being cut per second is increasing; the induced voltage is increasing.

From B to C. The induced voltage continues to increase until, at C, a maximum is reached. At this point, the conductor is moving perpendicular to the flux lines and cutting a maximum number per second.

From C to D. The number of flux lines being cut per second gradually decreases until the point D is reached. At D the voltage induced is the same as at B.

From D to E. The induced voltage continues to decrease until, at E, it is zero again. At this point, the conductor is again moving parallel to the flux lines.

From E to F. The conductor is now moving upward. The induced voltage is therefore in the opposite direction. It increases from zero until at F it is equal in magnitude and opposite to the induced voltage at B or D.



Graph of Induced Voltage

From F to G. The induced voltage increases until a maximum occurs at G; this maximum is in a direction opposite to that at C.

From G to H. The voltage decreases again until, at H, it is equal to that at F.

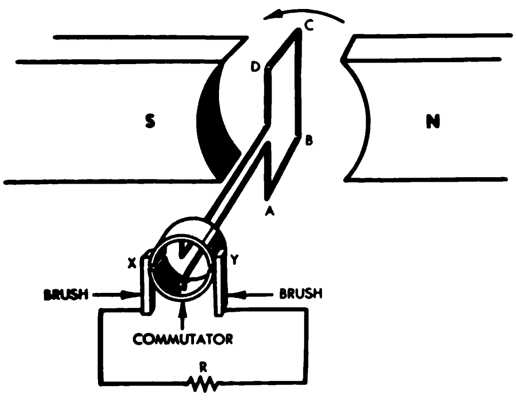
From H to A. The decrease continues until the conductor returns to its starting position at A where the induced voltage is again zero.

The entire process is repeated every time the conductor makes a complete revolution.

The Sine Wave

If you make a graph of the voltage induced in the conductor at the various points discussed, you will obtain an interesting curve. Plot the various points in degrees along the horizontal axis and the magnitude of the voltage induced at each point along the vertical axis. Use the horizontal axis as a reference for zero voltage. Consider the voltage induced in one direction as positive and that in the opposite direction as negative. The result obtained is shown in the illustration, Graph of Induced Voltage.

A Simple DC Generator

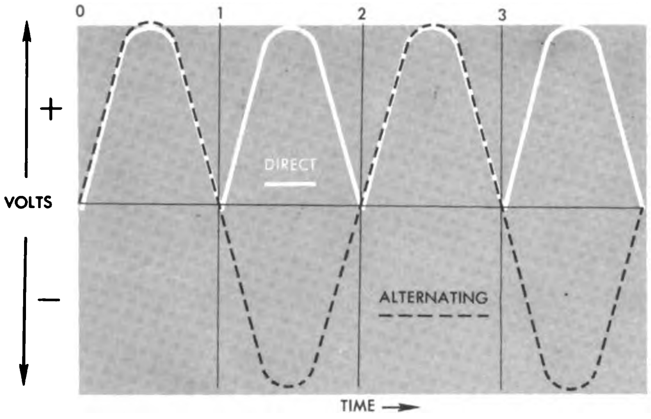


The curve is similar to one developed in mathematics for the variations of the function of an angle as the angle is increased. Since this function is called a *sine,* the curve is known as a *sine wave.* It is used to represent the voltage or current produced by a generator of the type discussed. Because this output periodically changes (alters) its direction, it is called *alternating current* or AC, and the generator is an *AC* generator.

A Simple DC Generator

The output of the AC generator is transferred to the external circuit by means of slip rings to brushes. Therefore, the AC voltage induced in the armature is AC when delivered to the external circuit. A device known as a *commutator* converts the AC voltage induced in the coil to DC voltage in the external circuit. A commutator is merely a divided ring made of segments insulated from one another. One brush makes contact with each segment of the commutator as shown in the illustration to the left, A simple DC Generator.

The brushes are so set with respect to the magnetic field that they change segments each time the current is passing through zero. Therefore, the current in the external circuit is always flowing in one direction, although its magnitude is continuously varying. A current of this type is known as *pulsating DC.* A graph of the pulsating DC current in the external circuit is represented by the solid line in the illustration on the next page, Output Voltage of a DC Generator. The dotted line represents the AC current in the coil. The pulsations — rises and falls — of the output voltage are known as *ripples.*



Output Voltage of a DC Generator

A DC generator always has a rotating armature and a stationary field. The rotating armature may consist of many coils; however, the commutator must have as many pairs of segments as the armature has coils. Since such coils are placed at different positions around the armature and connected in series, the resultant output never drops to zero as it does with one coil. Although increasing the number of coils and commutator segments reduces the ripple in the output immensely, it is impossible to remove the ripple completely.

Field Magnets

The magnetic field used in generators can be produced by permanent magnets or electromagnets. Permanent magnets are used in very small machines called magnetos. Most other generators have electromagnets which must be supplied with current. If the current is supplied by an outside source (a battery or another generator), the generator is *separately excited.* If the generator itself supplies current for the field magnets, it is called a *self-excited generator.* AC generators must be separately excited while DC generators are usually self-excited.

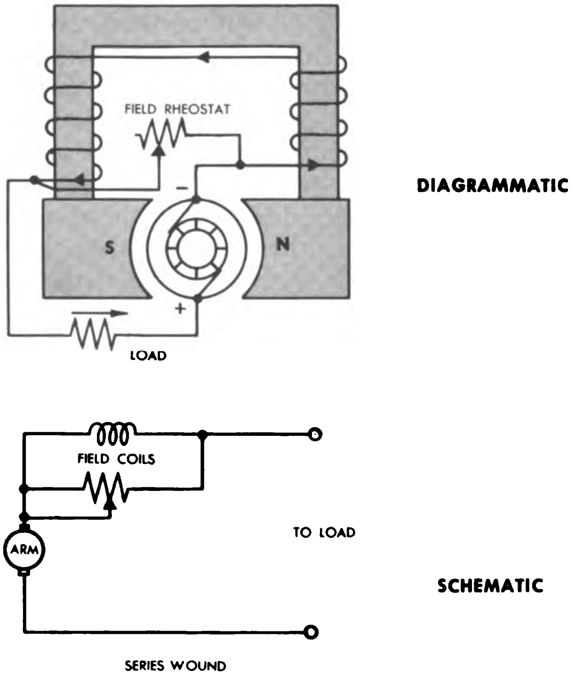
The field windings of a self-excited DC generator are connected in series or in parallel with the armature. When the generator is started, the armature begins to turn in the weak field which remains in the core of the electromagnets due to residual magnetism. An electromotive force is, therefore, induced in the coils of the armature which causes a current to flow in the field windings. This current makes the magnetic field stronger and increases the induced EMF. This, in turn, increases the field strength which again causes an increase in the induced EMF. The effect is cumulative until the generator reaches some fixed value of voltage output. The resistance of the field windings and the saturation of the core of the field magnets prevents this buildup from continuing indefinitely.

TYPES OF DC GENERATORS

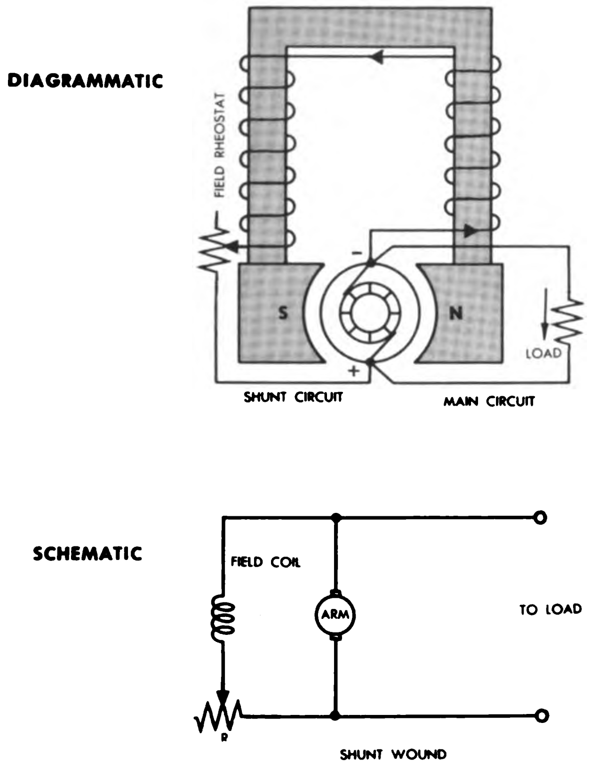
There are three types of DC generators — series-wound, shunt-wound, and shunt-series or compound wound. The difference in type is due to the relation of the field winding to the external circuit.

Series-wound Generators

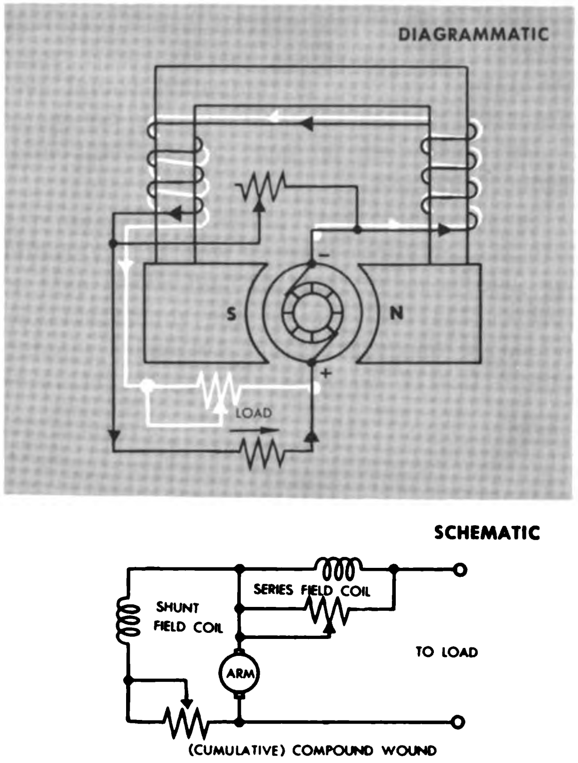
A series-wound generator has its field windings connected in series with the armature and the external circuit, called the load. The field coils consist of a few turns of low resistance wire because the load current is flowing through them. The method of wiring used for a series-wound generator is shown in the illustration below.



Series-wound Generator



Shunt-wound Generator



Compound-wound Generator

The ability of a generator to have a constant voltage output under varying load conditions is called its *voltage regulation.* Series-wound generators have very poor voltage regulation. The greater the load becomes the greater is the current through the field coils and the greater the induced EMF. Terminal voltage *increases with increasing load and decreases with decreasing load.* Because of their poor regulation, series-wound DC generators are not used frequently.

The output voltage of a series-wound generator may be controlled by a rheostat in parallel with the field windings as shown in the schematic.

Shunt-wound Generators

A shunt-wound generator is one in which the field windings are connected in parallel (shunt) with the armature and the load as shown in the illustration above. Since the field coils are in parallel with the load, the current through them is wasted as far as output is concerned. The field coils are, therefore, made of many turns of high resistance wire to keep the current flow through them small.

A shunt-wound generator is suitable if the load is constant; however, if the load fluctuates, the voltage varies also. This variation is due to the fact that the output voltage is equal to the induced voltage minus the IR drop across the armature. Since the armature supplies the current to the field and to the load, any increase in the load current increases the IR drop across the armature. This IR drop lowers the output voltage and the voltage across the field, thus reducing the field strength and further lowering the output voltage until some balance is reached. The terminal voltage of a shunt-wound generator *decreases with increasing load and increases with decreasing load.*

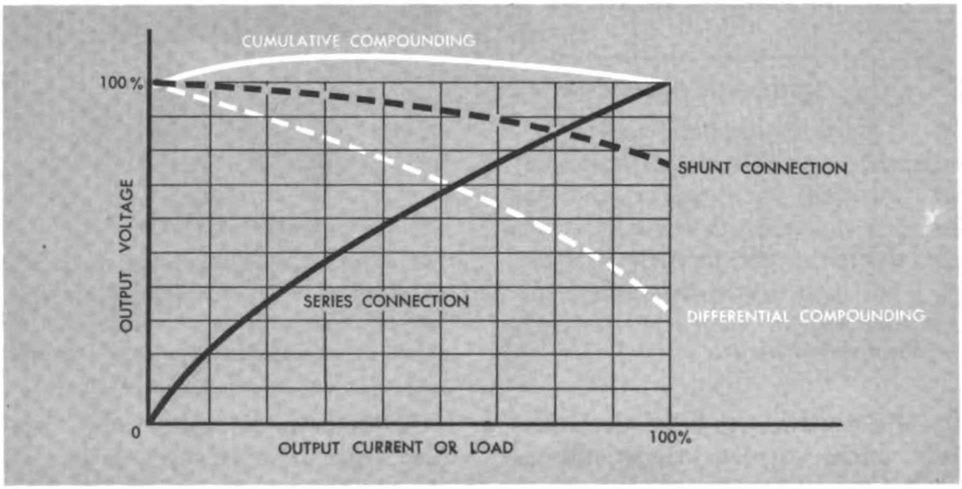
The output voltage of a shunt-wound generator may be controlled by means of a rheostat in series with the shunt field as shown in the schematic.

Compound-wound Generators

A compound-wound generator combines a series winding and a shunt winding in such a way that it uses the characteristics of each to advantage. The shunt field is normally the stronger of the two; the series field is used only to compensate for effects that tend to decrease the output voltage. A compound-wound generator and a schematic representation are shown in the illustration on page 59, A Compound-wound Generator.

The series field may vary from a fraction of a turn to a few turns of heavy wire wound around the field poles so that its field either aids or opposes the field flux of the shunt winding. When the magnetic field of the series winding aids the field of the shunt winding, the generator is *cumulative-compounded.* This is the type shown in the illustration and is more common. When the flux of the series winding opposes that of the shunt winding, the generator is *differential-compounded.* In this type, the series field tends to lower the output voltage as the load increases. Differential-compounded generators are sometimes used in welding generators where a machine must be able to withstand short-circuiting.

A summary of the characteristics of the four types of generators discussed is shown in the graph below.



Characteristics of DC Generators

Chapter 10

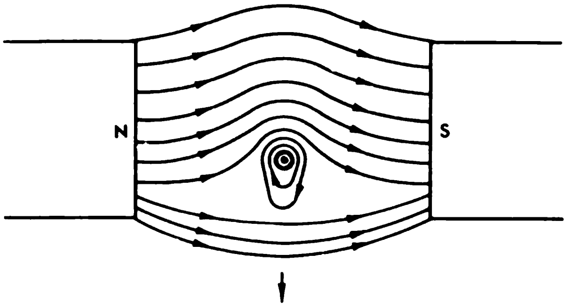
Direct Current Motors

Interchange the input and the output of a generator — supply electrical energy and obtain mechanical energy — and you have a motor. If the electrical energy supplied is DC, you have a DC motor. A DC motor is a rotating machine which transforms direct current energy into mechanical energy. A motor, like a generator, consists of two principal parts — a field assembly and an armature assembly. The armature is the rotating part in which current-carrying wires acted upon by the magnetic field are caused to rotate. The magnetic field may consist of a permanent magnet or electromagnets; however, electromagnets are used almost exclusively in DC motors, for they lend themselves to a wide variety of uses.

MOTOR PRINCIPLE

Before the motor is discussed, an examination of the principle of electromagnetism upon which its operation depends may be helpful. If you place a wire carrying current in a uniform magnetic field so that it is at right angles to the direction of flux, the wire moves. The direction of the movement is at right angles to both the magnetic field and to the wire. The cause of the movement is the force exerted by the interaction between the magnetic field around the wire and the magnetic field present.

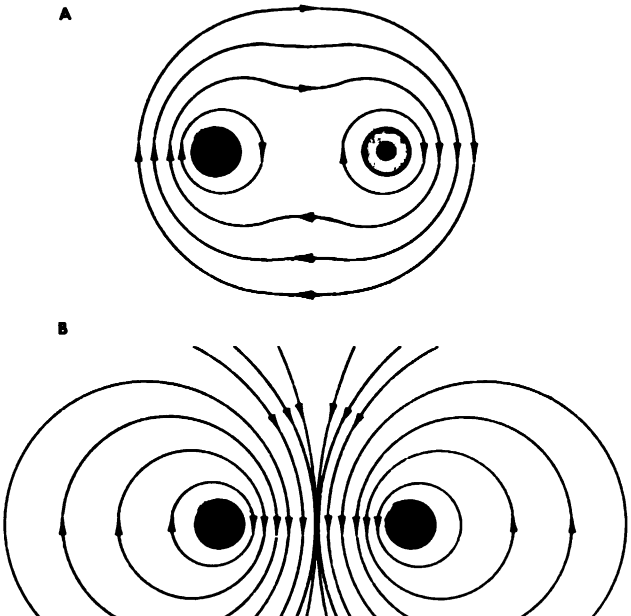
Examine the illustration shown at the top of page 62. Note that the lines of force of the magnetic field are from left to right and that electron flow in the wire is out of the page. By the left-hand rule, you find that the direction of the lines of force around the conductor is clockwise. Therefore, the two fields reinforce each other above, and oppose each other below the conductor, as shown. This results in a stronger field above the conductor and a weaker field below it. The conductor, therefore, is forced downward in the direction of the weaker field and away from the stronger one. If the direction of the magnetic field or of electron flow were reversed, the wire would be forced upward.



Field Surrounding a Conductor in a Magnetic Field

Force between Parallel Conductors

Two wires carrying current in the vicinity of one another exert a force on each other because of their magnetic fields. An end view of two conductors is shown in the illustration, Field Surrounding Parallel Conductors. In A, electron flow in both conductors is out of the page and the magnetic fields are clockwise around the conductors. Between the wires the fields cancel because the directions of the two fields oppose each other. The wires are forced in the direction of the weaker field — toward each other. This force is one of *attraction.*



Fields Surrounding Parallel Conductors

In B, the electron flow in the two wires is in opposite directions. The magnetic fields are therefore clockwise in one, and counterclockwise in the other, as shown. The fields reinforce each other between the wires and the wires are forced in the direction of the weaker field — away from each other. This force is one of *repulsion.*

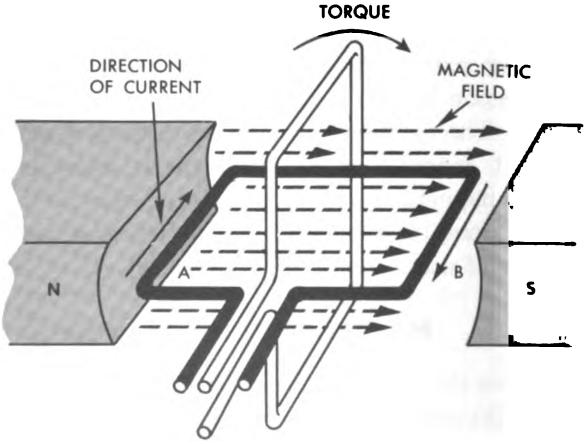
To summarize: Conductors carrying current in the same direction tend to be drawn together; conductors carrying current in opposite directions tend to be repelled from each other.

Developing a Torque

If a coil in which current is flowing is placed in a magnetic field, a force is produced which will cause the coil to rotate. This tendency of a force to produce rotation is called *torque.*

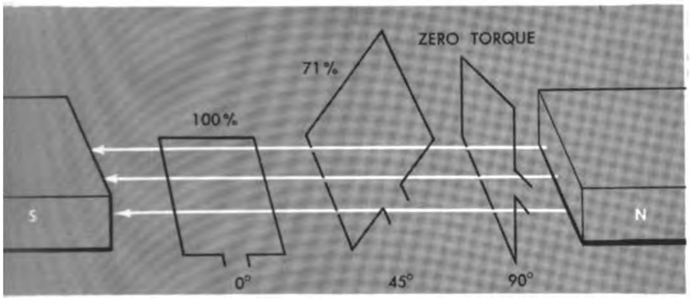
The illustration below shows a view of a coil in cross-section in a magnetic field. Current flows into the page at A and out of the page at B. The force on conductor A pushes it upward and that on conductor B pushes it downward. Since both forces are additive, their effect is to make the coil turn. A torque is produced which causes the coil to turn in a clockwise direction. The coil rotates until it reaches a position such that it is perpendicular to the flux lines. When it reaches this position, the force on each conductor is in the same direction and the coil stops rotating.

The amount of torque developed in a coil carrying current and placed in a uniform magnetic field depends upon the strength of the field, the number of turns in the coil, and the position



Developing a Torque

of the coil in the field. The torque is maximum when the plane of the coil is parallel to the flux lines and minimum when at right angles. It varies as a sine wave between these positions as shown in the illustration below.

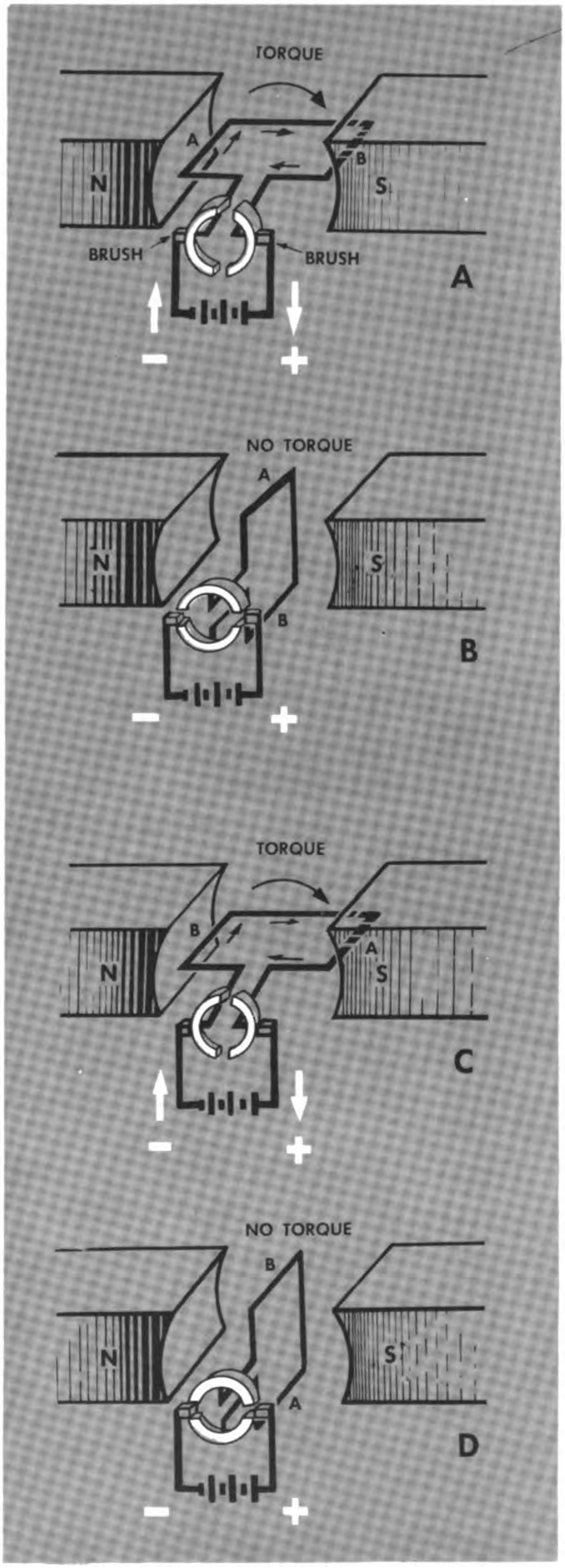


Variation in Torque Exerted on a Coil

THE BASIC MOTOR

The problem in the DC motor is to keep the coil rotating when it reaches the position where it is perpendicular to the field. This is accomplished by using a commutator which reverses the direction of electron flow each time the coil is about to stop. In the DC generator the commutator changes AC in the coil to DC in the output. The process is reversed in the motor; the DC input is changed to AC in the coil.

Assume the electron flow in the coil is as shown in the illustration, The Basic DC Motor (diagram A). In this position, the torque is maximum and the coil rotates in a clockwise direction. As the coil rotates, the torque becomes less and less until it becomes minimum when the coil is perpendicular to the field (diagram B). The momentum of the coil carries it a little beyond this minimum position. If the current were to continue to flow in the same direction, the coil would move back to the perpendicular position and stop. However, the commutator segments cause a reversal in the direction of current at this instant. The force on A is now downward and on B upward but the torque is weak. The rotation continues in a clockwise direction, and the torque builds up until it becomes maximum when the coil is parallel to the field again. However, conductor B is now on the left and A is on the right. (diagram C).



The Basic DC Motor

The clockwise rotation continues with the torque decreasing until it is minimum in the perpendicular position again (diagram D). Once again the change in direction of current flow by the commutator at this instant causes the clockwise rotation to continue. The torque builds up until the coil returns to the position shown in diagram A. This rotation continues as long as DC is applied, and the commutator reverses direction each time the torque becomes a minimum.

In the discussion of the one-coil basic motor, the commutator was considered as having two segments. In practice, however, the DC motor has a great many coils wound on the armature. In almost every position coils are near the poles of the magnet. The torque is, therefore, stronger and steadier. The commutator (as in the generator) consists of as many pairs of segments as there are coils on the armature. The field magnets are usually electromagnets whose current is supplied from the same DC source as the armature.

BACK EMF

Since the motor consists of an armature rotating in a magnetic field, a voltage is induced in the armature. Therefore, a motor is also a generator. Because the motion of the armature of the motor is caused by current flowing in it, the induced voltage causes a current to flow which opposes this current (Lenz's law). This induced voltage is therefore called *back* or *counter* EMF. The faster a motor rotates, the greater is this back EMF.

The current which flows through the armature windings depends upon the voltage applied, the back EMF, and the resistance of the wire. If the voltage applied to a motor were 110 volts and the counter EMF were 105 volts, the current flow through the armature windings would be the same as though 5 (110-105) volts were applied. In this case, the current flow would be 10 amperes if the resistance of the wire were 0.5 ohms.

THE SERIES MOTOR

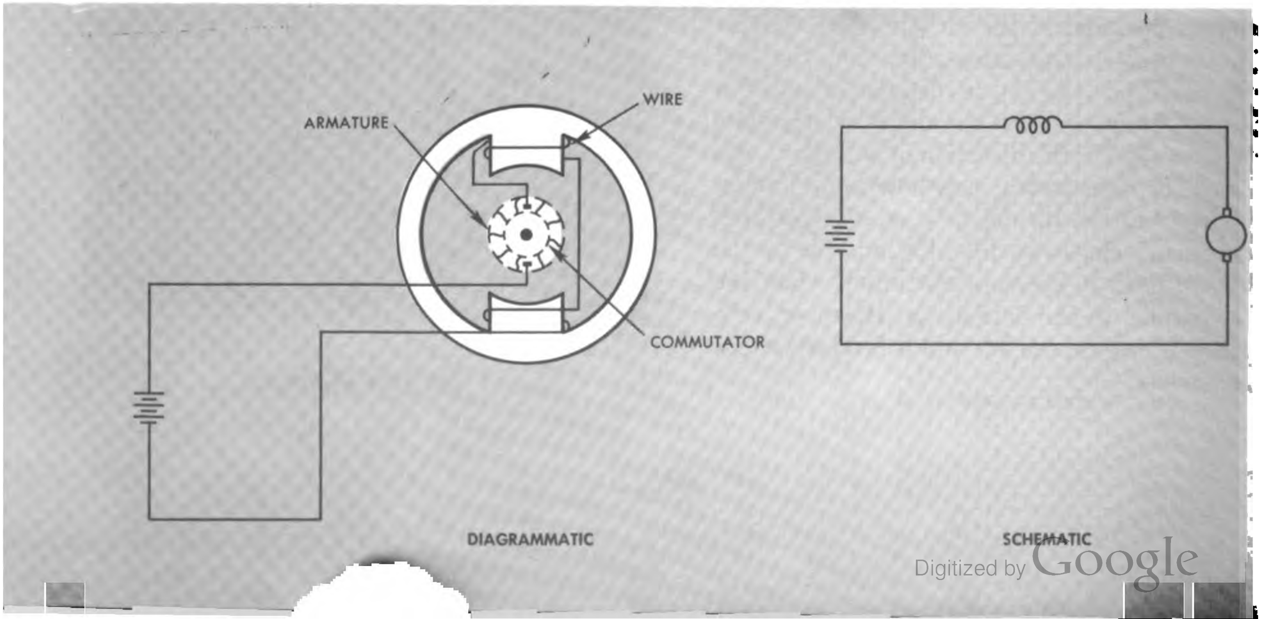
Motors, like generators, are classified according to the method in which their field coils and armature coils are connected. Thus, there are *series*, *shunt*, and *compound* motors.

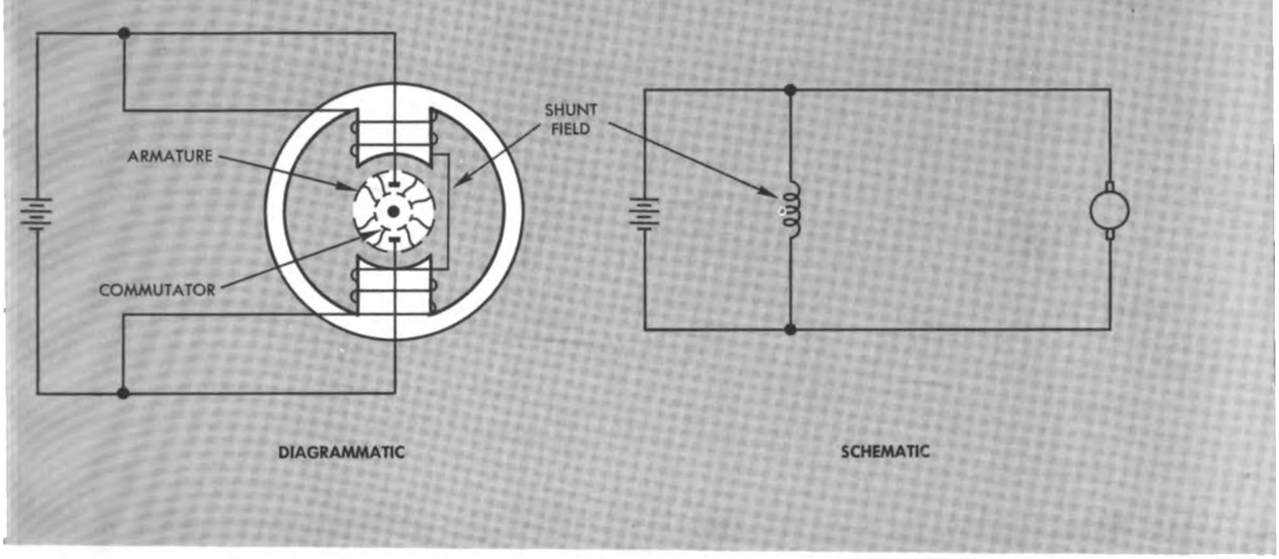
The field windings of a series motor consist of a few turns of heavy wire connected in series with the armature as shown in the illustration below. Since the same current flows through the armature and the field windings, an increase in current strengthens both magnetic fields. Because of the low resistance of the series field, the series motor draws a large current when starting. It, therefore, has a *high starting torque.*

The speed of a series motor depends upon the load. It runs at high speed with light load and at low speed with heavy load. If the load is removed entirely, the motor may attain a speed so high that the armature would fly apart.

The variation in the speed of a series motor with changing load may be explained in terms of back EMF. DC motors accelerate until the back EMF reduces the current to such an extent that the power input is just sufficient to supply the

A Series Motor





A Shunt Motor

losses of the motor and allow it to accomplish its work. In the case of a series motor with no load, the back EMF must be almost as large as the line voltage. If the current is small, the magnetic field is weak and the armature must travel faster to build up the necessary back EMF. With no load, the current is so small that the speed builds up until the armature flies apart. A series motor must, therefore, *never* be started without a load nor be connected to a load by means of a belt which might slip off or break.

Series motors have many applications where high starting torque is needed under heavy load conditions. A trolley car may use a series motor to advantage since it has great mass and needs high starting torque. Series motors are often used in aircraft as starters and for raising and lowering landing gears, cowl flaps, and wing flaps.

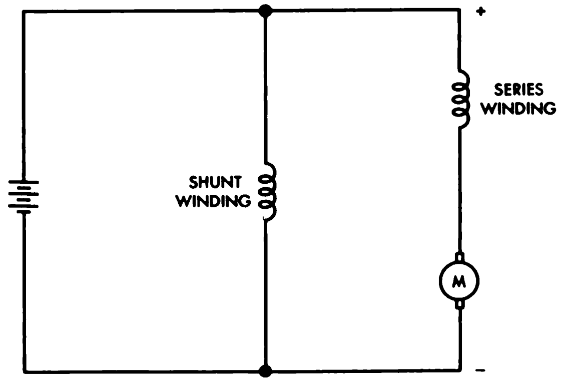
THE SHUNT MOTOR

The field winding of a shunt motor is connected in parallel with the armature windings as shown in the illustration above. The field winding is made up of many turns of fine wire and connected directly across the power supply. The current through the field is fairly constant and the torque varies with variations in armature current only.

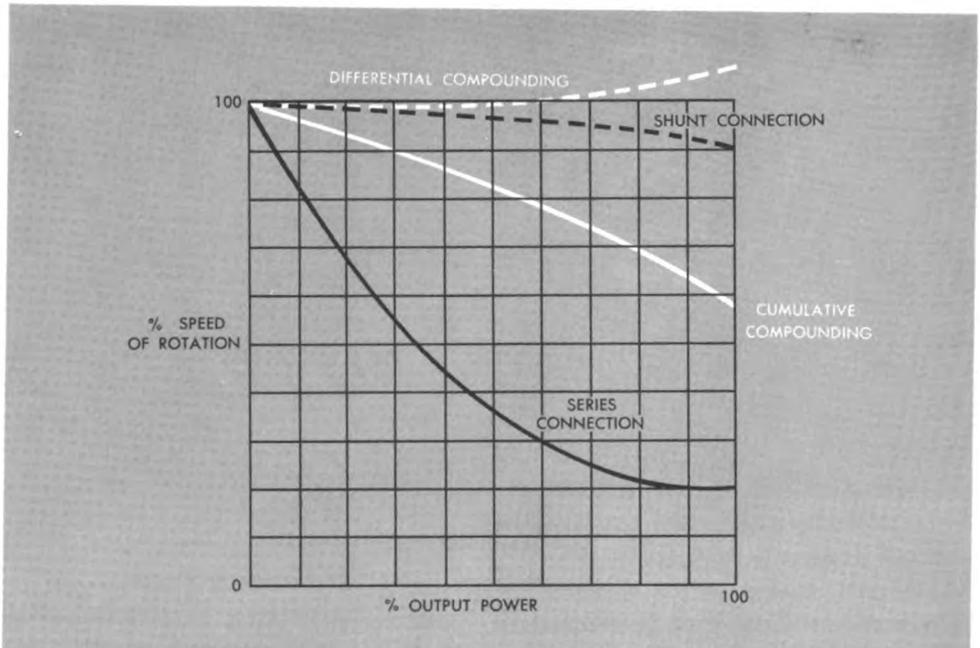
The speed of a shunt motor reaches a steady state before becoming excessive. With no load, the speed of the motor will be only slightly faster than under load. A shunt motor has an almost *constant speed* but a starting torque which is less than that of a series motor of the same size.

THE COMPOUND MOTOR

A compound motor is a combination of a series and shunt motor. It has a field with two windings — a series winding and a shunt winding — as shown in the schematic illustration below. The starting torque of a compound motor is lower than that of a series motor but higher than that of a shunt motor. Changes in load affect the speed of a compound motor less than in a series motor but slightly greater than in a shunt motor. The compound motor is very common in aircraft where a constant speed and high starting torque are needed.



Schematic of a Compound Motor



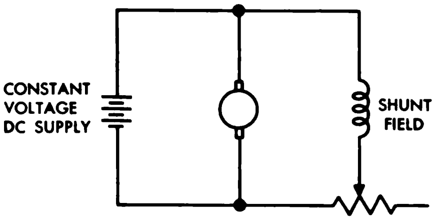
Effect of Change of Load on Speed of a DC Motor

A graph of the variations in speed with changes of load of the various types of DC motors is shown above. Note that both a cumulative-compound and a differential-compound motor are graphed. The cumulative-compound motor is the more common in aircraft.

MOTOR SPEED

A motor whose speed can be controlled is called a variable-speed motor. The speed of a motor may be altered by *varying the current in the field winding* or *by varying the armature voltage*. When the field current is decreased, the flux is decreased and the counter EMF is decreased. The motor therefore speeds up. An increase in field current, on the other hand, has the effect of slowing down the motor. If the voltage across the armature is decreased, the motor speed decreases; an increase in voltage across the armature increases the speed of the motor.

In the shunt motor, speed is usually controlled by a rheostat in series with the field windings as shown in the illustration to the right. When the rheostat resistance is increased, the current through the field winding is decreased. The decreased flux momentarily decreases the counter EMF. The motor then speeds up to increase the counter EMF and keep the armature current at the same value. In similar fashion, a decrease in rheostat resistance increases the current flow through the field windings and causes a reduction in motor speed.



Shunt Motor with Speed Control

In a series motor, the rheostat speed control is connected either in parallel or in series with the armature windings. Moving the rheostat in a direction which lowers the voltage across the armature slows the motor down; moving it in a direction which raises the voltage across the armature speeds the motor up.

Reversing Direction

To reverse the direction of rotation of a motor, reverse either the field or armature connections, *not both*. If the connections to an external source are reversed, the direction of rotation is not reversed because both field and armature connections are changed together.

One method of reversing the direction of rotation is shown below in the illustration, A Split Field Motor. This is a series motor with a split field winding. When the switch is placed in position A, electrons flow through the lower portion of the winding creating a magnetic field with a north pole at the bottom and a south pole at the top. The motor rotates in one direction. In position B, electrons flow through the upper portion of the winding creating a magnetic field with a north pole at the top and a south pole at the bottom. The direction of flux is reversed, but the direction of electron flow through the armature remains the same. The direction of rotation is therefore reversed.

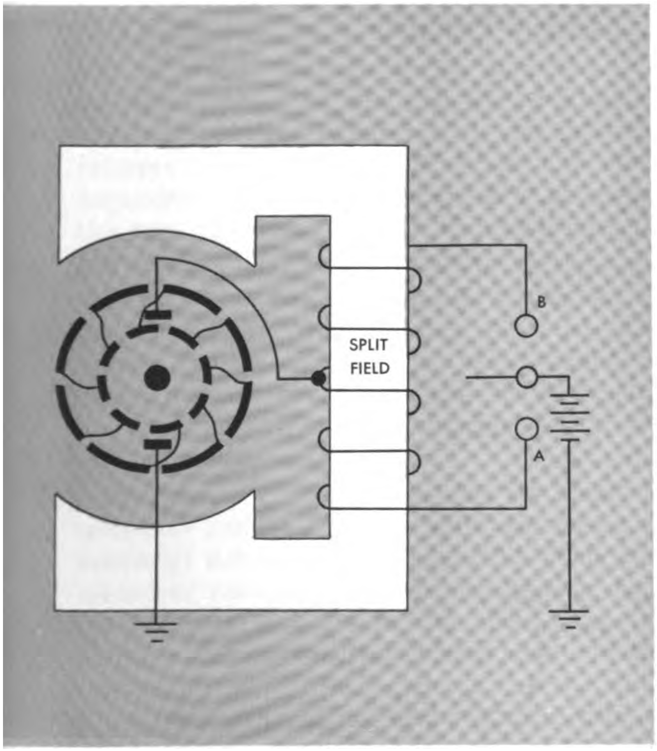
Another method of reversing the direction of rotation is shown in the illustration, Reversing Direction with Switch. A double-pole, double-throw switch is used to reverse the direction of current through the field but not through the armature. With the switch in position AA, electron flow through the field winding is from right to left and in position BB, from left to right.

LOSSES IN DC MACHINES

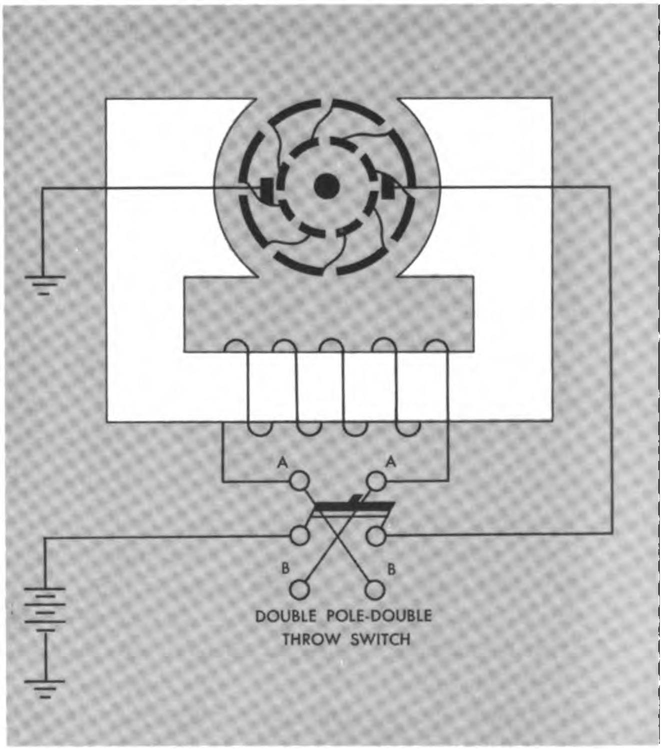
Losses occur whenever electrical energy is converted to mechanical energy (in the motor) or mechanical energy is converted to electrical energy (in the generator). For the machine to be efficient, these losses must be kept to a minimum. Some losses are *electrical* and some are *mechanical.* Electrical losses are classified as copper losses and iron losses; mechanical losses occur in overcoming the friction of various parts of the machine.

Copper Losses

Copper losses are the losses which occur in forcing electrons through the copper windings of the armature and the field. These losses are proportional to the square of the current. They are sometimes called I^2R losses because they are due to the power dissipated in the form of heat in the resistance of the field and armature windings.



A Split Field Series Motor



Reversing Direction with Switch

Iron Losses

Iron losses are subdivided into *hysteresis* and *eddy current* losses. Hysteresis losses result because the *armature* is revolving in an alternating magnetic field. It, therefore, becomes magnetized first in one direction and then in the other. The residual magnetism of the iron or steel of which the armature is made causes these losses. Since the field magnets are always magnetized in one direction (DC field), they have no hysteresis losses.

Eddy current losses occur because the iron core of the armature is a conductor revolving in a magnetic field. This sets up an EMF across portions of the core causing currents to flow within the core. These currents heat the core and may cause damage to the windings if they become excessive. As far as the output is concerned, the power consumed by eddy currents is a loss. To reduce eddy currents to a minimum, a *laminated* core is usually used. A laminated core is one made of thin sheets of iron electrically insulated from each other. The insulation between laminations reduces eddy currents because it is "transverse" to the direction in which these currents tend to flow. However, it has no effect on the magnetic circuit. The thinner the laminations, the more effectively does this method reduce eddy current losses.

SUMMARY

As an Air Force observer, you must be familiar with the fundamental concepts of electricity, for modern radio and radar equipment is based on electronics. So far we have dealt with the basic principles and laws of electricity and how they govern direct current. We have seen how some of these various principles and laws are applied in circuits, generators, and motors.

Direct current circuit problems were illustrated in order to familiarize you with the procedure for obtaining current, resistance, or voltage in direct current circuits.

Storage batteries and dry batteries supply direct current through chemical action, while generators and motors are electro-mechnical devices that depend on magnetic fields for their operation.

Electricity and magnetism are interrelated subjects. An electric current creates a magnetic field, and a moving magnetic field induces a voltage in a conductor.

Chapter 11

Meters

Meters are used in repairing, maintaining, and troubleshooting electronic equipment. As an observer, you will find that the operation of electronic devices requires that you properly interpret meter readings. The best and most expensive measuring instrument is of no use to the man who does not know what he is measuring or what the readings indicate.

The theory, construction, and limitations of the meters you will use most often are discussed in this chapter. An understanding of this important information will enable you to use meters properly and to interpret their readings correctly.

In chapter 4, you learned that the ammeter measures current flow, the voltmeter measures potential difference or voltage, the ohmmeter measures resistance, and the multimeter is a combination of these meters. You must remember that the function of a meter is to measure quantities. When a meter is inserted into a circuit, it should not change the conditions of the circuit. This requirement is not easily achieved because most meters require current for their operation, and this current must come from the circuit being measured. One of the objectives of this chapter will be to show you how, and when, your meter will not indicate the actual circuit conditions.

EFFECTS OF CURRENT

Since it is not possible to count the invisible electrons as they move through a conductor, meters measure current by the effects these electrons produce on measurable quantities. The effect of current may be classified as follows: chemical, physiological, photoelectric, piezoelectric, thermal, and electromagnetic.

Chemical

When an electric current is passed through certain solutions, a chemical reaction takes place and a deposit forms on one electrode. The amount of this deposit is proportional to the amount of current. You are acquainted with this chemical effect of current from your study of the battery. Industrially, it is useful in electroplating and electrolysis. Although the chemical effect is useful in defining the standard ampere (the amount of current which causes

.001118 grams of silver to be deposited in one second from a 15% solution of silver nitrate), it is of no practical use in meters.

Physiological

The reaction of the human body to an electric current is what is meant by the physiological effect of current. An electric shock, although painful at times, is too difficult to evaluate quantitatively and is, therefore, not practical for use in meters.

Photoelectric

When electrons strike certain materials, a glow appears at the point of contact. The picture tube of a television set and the scope of a radar set illustrate this effect. Using the intensity of the light produced as a means of measuring the amount of current is neither accurate nor practical.

Piezoelectric

Certain crystals such as quartz and Rochelle salt become deformed when a voltage is impressed across two faces. This effect is not visible to the human eye and is, therefore, impractical for use in meters.

Thermal

When a current flows through a resistance, heat is produced. The amount of heat produced is equal to I^2R. From this relationship you can see that heat varies as the square of the current. Meters which employ the thermal effect in their operation are common.

Electromagnetic

Whenever electrons flow through a conductor, a magnetic field which is proportional to the current is created. This effect is useful for measuring current and is employed in many practical meters.

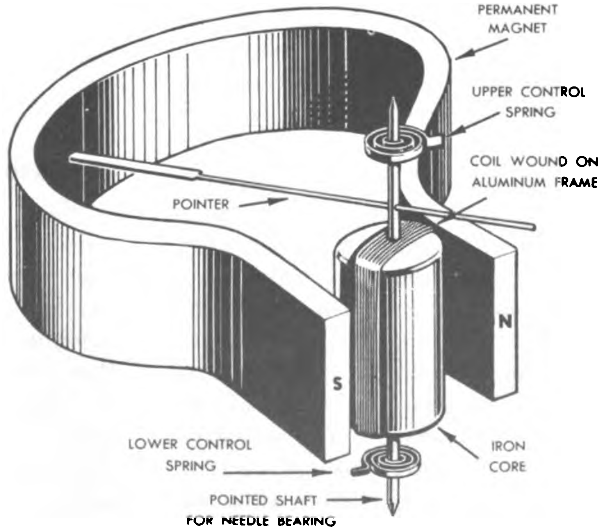
You can see from the discussion that the first four effects are of no practical importance to the observer. The last two effects, thermal and magnetic, are of practical use in meters. Since most of the meters used by the observer have D'Arsonval movements which operate because of the magnetic effect, only this type will be discussed in detail.

THE D'ARSONVAL METER

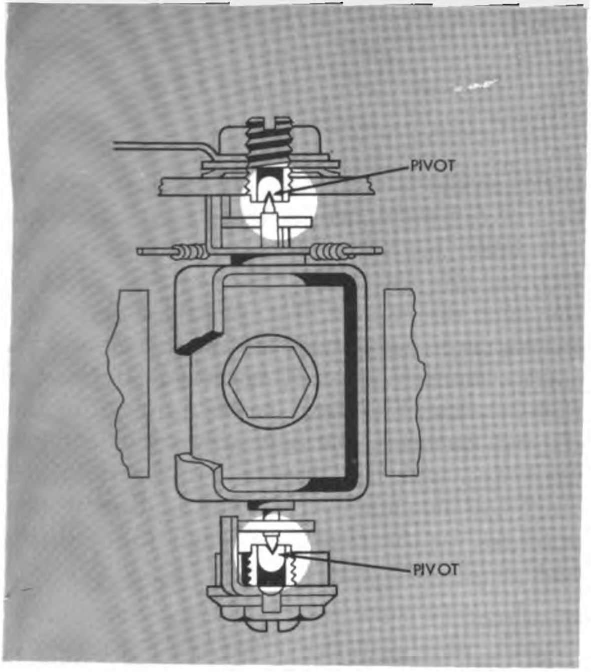
The D'Arsonval meter is an instrument designed for measuring direct current flowing in an electrical circuit and consists of the following parts: a permanent magnet, a moving element with attached pointer, a moving element mounting, bearings, and a case which includes terminals, a dial, and screws. Each part and its function are described in the discussion which follows.

The *permanent magnet* furnishes a magnetic field which will react with the magnetic field set up by the moving element.

The *moving element* is mounted so that it is free to rotate when energized by the current to be measured. A pointer which moves across a calibrated scale is attached to this element which is a *moving coil* in the D'Arsonval. A



Moving Element with Pointer and Springs

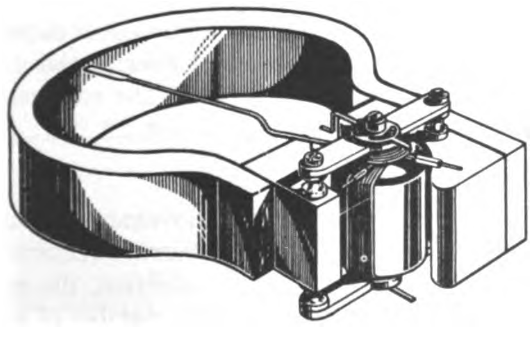


Method of Mounting Moving Elements

moving coil mechanism is shown in the illustration on page 70.

The *controlling element* is a spring, or springs, whose main function is to provide a counter or restoring force. The strength of this force increases with the turning of the moving element and brings the pointer to rest at some point on the scale. Two springs are generally used; they are wound in opposite directions to compensate for the expansion and contraction of the spring material due to temperature variation. The springs are made of nonmagnetic material and conduct current to and from the moving coil in some meters.

The *moving-element mounting* consists of a shaft with very hard pivot points to carry the moving coil or other movable element. The pivot points are so fitted into highly polished jewels or very hard glass bearings that the moving element can rotate with very little friction. Another type of mounting has been designed in which the pivot points are reversed and the bearings are inside the moving-coil assembly. A method of mounting moving elements is shown in the illustration to the left.



D'Arsonval Meter Movement

The *bearings* are highly polished jewels such as sapphires, synthetic jewels, or very hard glass. These are usually round in shape and have a conical depression in which the pivots rotate. They are set in threaded nuts which allow adjustment. The radius of the depression in the jewel is greater than the radius of the pivot point. This limits the area of contact surfaces and provides a bearing which, when operated dry, probably has the lowest constant friction value of any known type of bearing.

The *case* houses the instrument movement and protects it from mechanical injury and exposure. It also has a window for viewing the movement of the pointer across a calibrated scale.

The *terminals* are made of materials having very low electrical resistance. Their function is to conduct the required current into and away from the meter.

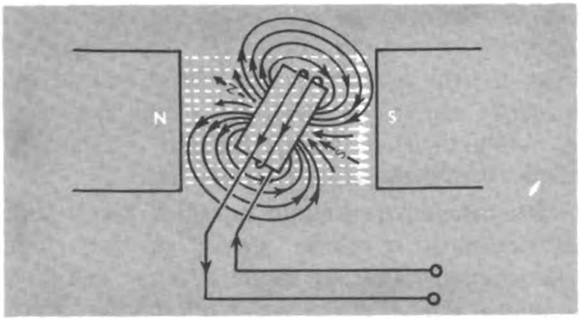
The *dial* has printed on it pertinent information such as the scale, units of measure, and meter uses.

HOW THE METER MOVEMENT WORKS

In the illustration to the left you can see how the major units are mounted in relationship to one another. Note that the coil portion of the moving element is in the magnetic field of the permanent magnet.

In order to understand how the meter works, assume that the coil of the moving element is placed in a magnetic field as shown in the illustration on page 72.

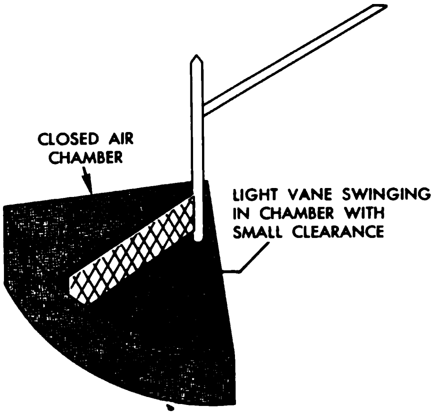
The coil is pivoted so that it is able to rotate back and forth within the magnetic field set up by the magnet. When the coil is connected in a circuit, current flows through the coil in the direction indicated by the arrows and sets up a magnetic field within the coil. This field has the same polarity as the adjacent poles of the magnet. The interaction of the two fields causes the coil to rotate to a position so that the two magnetic fields are aligned. This force of rotation (torque) is proportional to the interaction between the like poles of the coil



Effect of a Coil Placed in a Magnetic Field

and the magnet, and therefore to the amount of current flow in the coil. As a result, a pointer attached to the coil will indicate the amount of current flowing in the circuit as it moves across a graduated scale.

In the arrangement just discussed note that any torque sufficient to overcome the inertia and friction of moving parts causes the coil to rotate until the fields align. This uncontrolled movement would cause inaccurate current readings. Therefore, the turning motion of the coil is opposed by two springs. The value of the current flowing through the coil determines the turning force of the coil. When the turning force is equal to the opposition of the springs, the coil stops moving and the pointer indicates the current reading on a calibrated scale. The springs are made of conducting material and serve to conduct current to and from the coil. The pole pieces of the magnet form a circular air gap within which the coil is pivoted.



Air Damping

To obtain a clockwise rotation, the north pole of the permanent magnet and that of the coil must be adjacent. The current flowing through the coil must therefore always be in the same direction. The D'Arsonval movement can be used only for *DC measurements*, and the correct polarity must be observed. If the current is allowed to flow in the wrong direction through the coil, the coil will rotate counterclockwise, and the pointer will be damaged. Since the movement of the coil is directly proportional to the current through the coil, the scale is normally a linear scale.

Damping

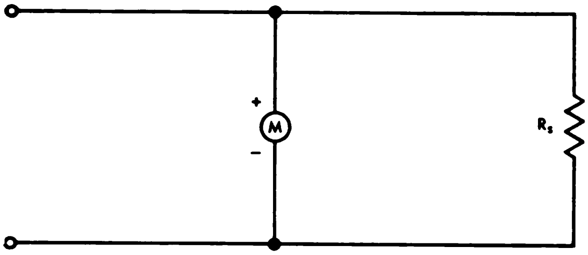
So that meter readings can be made quickly and accurately, it is desirable that the moving pointer overshoot its proper position only a small amount and come to rest after not more than one or two small oscillations. The term damping is applied to methods used to bring the pointer of an electrical meter to rest after it has been set in motion. Damping may be accomplished by electrical means, by mechanical means, or by a combination of both.

ELECTRICAL DAMPING. A common method of obtaining damping by electrical means is to wind the moving coil on an aluminum frame. As the coil moves in the field of the permanent magnet, eddy currents are set up in the aluminum frame. The magnetic field produced by the eddy currents opposes the motion of the coil. The pointer will therefore swing more slowly to its proper position and come to rest quickly with very little oscillation.

AIR DAMPING. Air damping is a common method of obtaining damping by mechanical means. As shown in the illustration to the left, a vane is attached to the shaft of the moving element and enclosed in an air chamber. The movement of the shaft is retarded because of the resistance which the air offers to the vane. Effective damping is achieved if the vane nearly touches the walls of the chamber.

Meter Sensitivity

The *sensitivity* of a meter movement is usually expressed as the amount of current required to give full scale deflection. In addition, the sensitivity may be expressed as the number of millivolts across the meter when full-scale current flows through it. This voltage drop is obtained



Meter Movement with Shunt

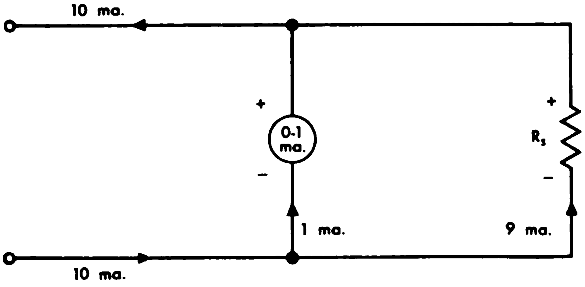
by multiplying the full-scale current by the resistance of the meter movement. A meter movement, whose resistance is 50 ohms and which requires 1 ma. for full scale reading, may be described as a 50-millivolt, 0-1 milliammeter.

EXTENDING THE RANGE OF AN AMMETER

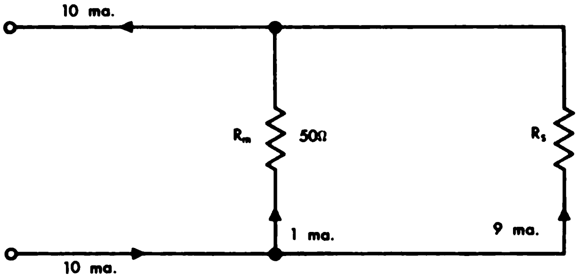
A 0-1 ma. meter movement may be used to measure currents greater than 1 ma. by connecting a resistor in parallel with the movement. The parallel resistor is called a *shunt* because it by-passes a portion of the current around the movement, extending the range of the ammeter. A schematic drawing of a meter movement, with a shunt connected across it to extend its range, is shown above.

Finding the Value of a Shunt

By applying the basic rules for parallel circuits, you can easily compute the value of a shunt resistor needed to extend the range of an ammeter. Suppose that you have available a 50 millivolt 0-1 ma. meter and find it necessary to measure current up to 10 ma. How would you do this? The first step is to draw a schematic of your meter shunted by a resistor, R_S, and list on it the available information. This schematic is shown below.



You know the sensitivity of the meter and can compute its resistance. You can establish how much current will flow in the branches since only a maximum of 1 ma. can flow through the meter. The equivalent circuit can then be redrawn as shown in the illustration below.



The voltage drop across R_S can be computed since it is the same as that across the meter, R_M.

$$\begin{aligned} E &= IR \\ &= .001 \times 50 \\ &= .050 \text{ volt} \end{aligned}$$

Once the voltage drop is known, R_S can be found by applying Ohm's law

$$\begin{aligned} R_S &= \frac{E_{R_S}}{I_{R_S}} \\ &= \frac{.050}{.009} \\ &= 5.55 \text{ ohms} \end{aligned}$$

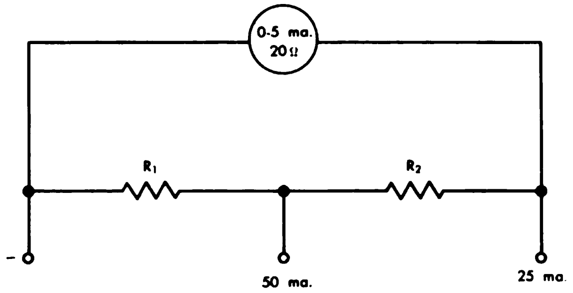
Note that the value of the shunt resistor is very small; however, this value is very critical. Resistors used as shunts for meters must have close tolerances, usually 1%.

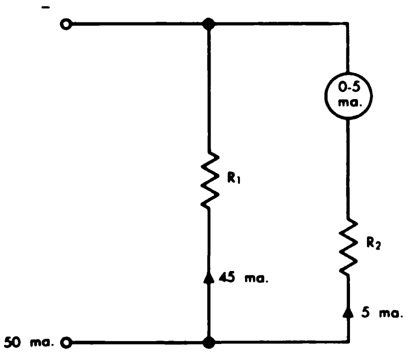
The Universal Shunt

The schematic drawing on page 74, the Universal Shunt, shows an arrangement whereby two or more ranges are provided by tapping the shunt resistor at the proper points. In this arrangement, a 0-5 ma. movement with a resistance of 20 ohms is shunted to provide a 0-25 ma. range and a 0-50 ma. range. The method of computing the value of the shunt resistor needed and the point at which it is to be tapped is as follows:

$$R_S = \frac{R_M\, I_M}{I_S}$$

Since the entire shunt is used on the 25 ma. range and 5 ma. flows through the meter, $I_S = 20$ ma. Therefore,





The Universal Shunt

$$R_S = \frac{20 \times .005}{.02} = 5 \text{ ohms}$$

On the 50 ma. range, R_1 is the shunt and R_2 is in series with the meter movement as shown in the second illustration. Since 5 ma. flows through R_2 and the meter movement, 45 ma. flows through R_1. The voltage across R_1 is equal to the voltage across R_2 and the meter. Therefore,

$$E_{R_1} = (.005)\ (20 + R_2)$$

$$\text{But, } R_1 = \frac{E_{R_1}}{.045}$$

Substituting for E_{R_1}, you have

$$R_1 = \frac{(20+R_2)\ (.005)}{.045}$$

$$\text{However, } R_1 = R_S - R_2$$

$$= 5 - R_2$$

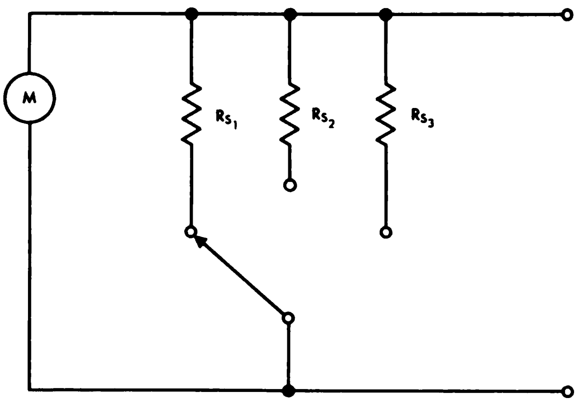
$$\text{Therefore, } (5 - R_2)\ (.045) = (20 + R_2)\ (.005)$$

$$.225 - .045R_2 = .1 + .005R_2$$

$$.050R_2 = .125$$

$$R_2 = 2.5 \text{ ohms}$$

$$\text{and} \quad R_1 = 2.5 \text{ ohms}$$



Variation of Universal Shunt

A variation of the universal shunt is shown in the illustration above. Here a switching arrangement is used which places the proper value of shunt needed for the desired range in the circuit. This is a common method used in many meters for extending the range of an ammeter.

In actual practice it is not usually necessary to compute the values of the various shunts needed because they are part of the meter and are connected internally. However, in some situations, an external shunt may be necessary to extend the range of a meter. In such cases, the computations are made as described.

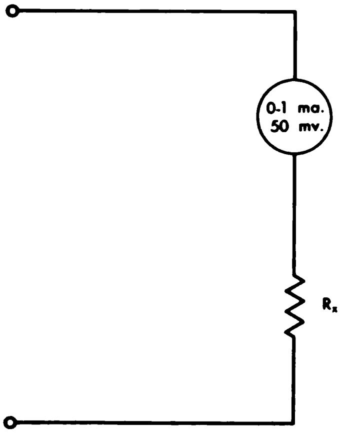
VOLTMETERS

A meter movement which measures current also has a voltage drop across it. Thus, the meter movement mentioned earlier was described as a 0-1 ma, 50 millivolt movement. A voltage drop of 50 millivolts would be of limited use for voltage measurements. It is therefore necessary to extend the range of the movement so that it will be of more practical use in measuring voltages.

Extending Voltmeter Range

To extend the voltage range of a meter movement, place a resistor in series with it as shown in the illustration at the top of page 75. This resistor (R_X) in series is called a multiplier.

The method of computing the value of a multiplier is to divide the voltage range desired by the full-scale current of the meter movement. Subtract the resistance of the meter movement from the result and the value of the multiplier is obtained. For example, if the meter movement



Computing the Value of a Multiplier

is a 50 millivolt, 0-1 ma. movement, the multiplier for a 10-volt range is obtained as follows:

$$R_t = \frac{10}{.001} = 10{,}000 \text{ ohms}$$

$$R_x = 10{,}000 - 50 = 9{,}950 \text{ ohms}$$

The illustration at A below shows a schematic of a voltmeter with three ranges: 0-10 volts, 0-100 volts, and 0-500 volts. If the meter has a 0-1 ma. 75 millivolt movement, the multipliers are calculated as follows:

$$R_1 + 75 = \frac{10}{.001} = 10{,}000 \text{ ohms}$$

$$R_1 = 9{,}925 \text{ ohms}$$

$$R_2 + 75 = \frac{100}{.001} = 100{,}000 \text{ ohms}$$

$$R_2 = 99{,}925 \text{ ohms}$$

$$R_3 + 75 = \frac{500}{.001} = 500{,}000 \text{ ohms}$$

$$R_3 = 499{,}925$$

The same meter movement may be connected in the manner shown at B to obtain a 10-volt, a 50-volt, and a 100-volt range. Note that R_1 is in series with the meter movement on the 10-volt range, R_2 is added in series on the 50-volt range, and R_3 is added in series on the 100-volt range. The values of R_1, R_2, and R_3 are found as follows:

$$R_1 + 75 = \frac{10}{.001} = 10{,}000 \text{ ohms}$$

$$R_1 = 9{,}925 \text{ ohms}$$

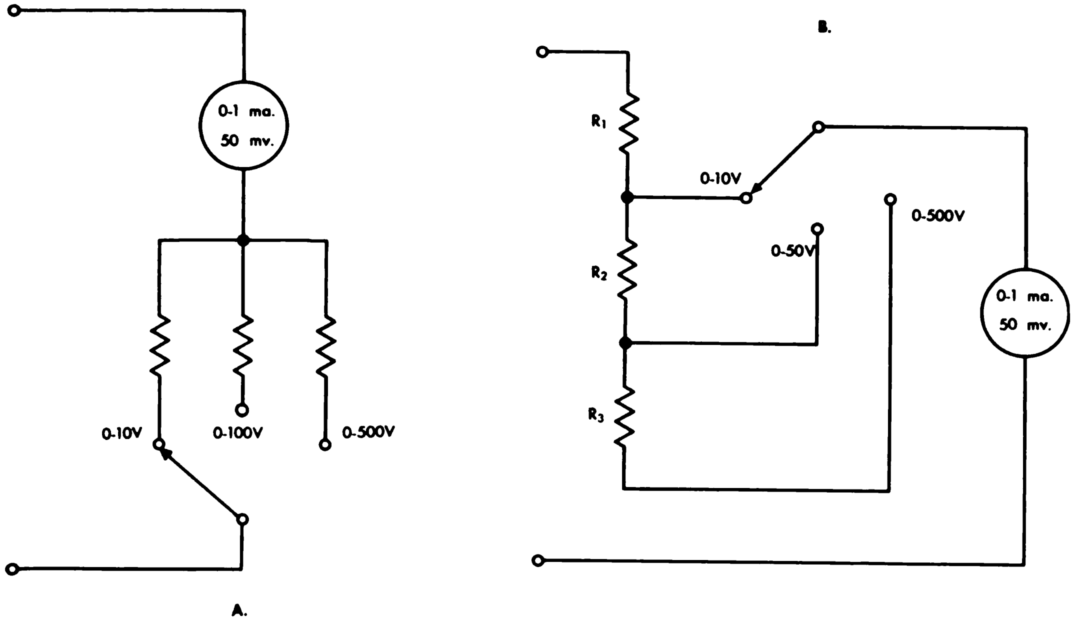
$$R_2 + 10{,}000 = \frac{50}{.001} = 50{,}000 \text{ ohms}$$

$$R_2 = 40{,}000 \text{ ohms}$$

$$R_3 + 50{,}000 = \frac{100}{.001} = 100{,}000 \text{ ohms}$$

$$R_3 = 50{,}000 \text{ ohms}$$

Schematics of Voltmeter with Three Ranges



Sensitivity of Voltmeters

The *sensitivity* of a voltmeter is expressed in *ohms-per-volt*. This value is obtained by dividing the total resistance in the meter circuit by the full-scale reading in volts. In each of the illustrated problems, the sensitivity is 1000 ohms-per-volt. This value is obtained as follows:

On the 10-volt range,

$$\frac{10{,}000}{10} = 1000 \text{ ohms-per-volt}$$

On the 50-volt range,

$$\frac{50{,}000}{50} = 1000 \text{ ohms-per-volt}$$

On the 100-volt range,

$$\frac{100{,}000}{100} = 1000 \text{ ohms-per-volt}$$

On the 500-volt range,

$$\frac{500{,}000}{500} = 1000 \text{ ohms-per-volt}$$

The sensitivity of a voltmeter is an indication of how accurately it measures voltages in a circuit. For most cases, a sensitivity of 1000-ohms-per-volt is satisfactory; however, if the circuit in which voltage is being measured has high resistance, a greater sensitivity is required for accuracy.

OHMMETERS

A basic meter movement combined with a source of EMF and resistors may also be used to measure resistance. Ohmmeters may be of the *series* or *shunt* type.

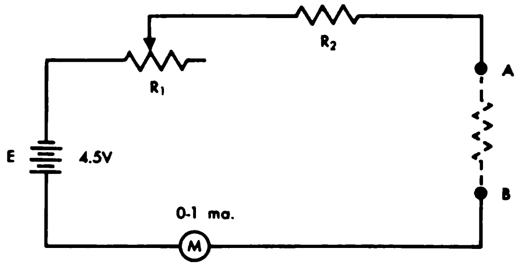
Series-Type Ohmmeter

A simplified schematic of an ohmmeter is shown to the right. E is a source of EMF; R_1 is a variable resistor used to zero the meter; R_2 is a fixed resistor used to limit the current in the meter movement; and A and B are test terminals across which the resistance to be measured is placed.

If A and B are connected together (short-circuited), the meter, the battery, and resistors R_1 and R_2 form a simple series circuit. With R_1 set so that the total resistance in the circuit is 4500 ohms, the current through the meter is 1 ma. and the needle deflects full scale. Since there is no resistance between A and B, this position of the needle is labeled zero. If a resistance equal to 4500 ohms is placed between terminals A and B, the total resistance is 9000 ohms and the current is .5 ma. This causes the needle to deflect half scale. This half scale reading is labeled 4500 ohms. Note that the half scale reading is equal to the internal resistance of the meter.

If a resistance of 9000 ohms is placed between terminals A and B, the needle deflects one-third scale. If terminals A and B are not connected (open-circuited), no current flows and the needle does not move. The left side of the scale is therefore labeled infinity to indicate an infinite resistance. Resistances of 13.5 K and 1.5 K placed between terminals A and B will cause a deflection of one-fourth and three-fourths of the scale, respectively.

Examine the ohmmeter scale shown in the illustration on page 77 to the right. Note that the scale is not linear and is crowded at the high resistance end. For this reason, it is good practice to use an ohmmeter range in which the readings are not too far from mid-scale. A good rule is to use a range in which the reading obtained does not exceed ten times, or is not less than one-tenth, the mid-scale reading. The useful range of the scale shown is, by this rule, from 450 ohms to 45,000 ohms.

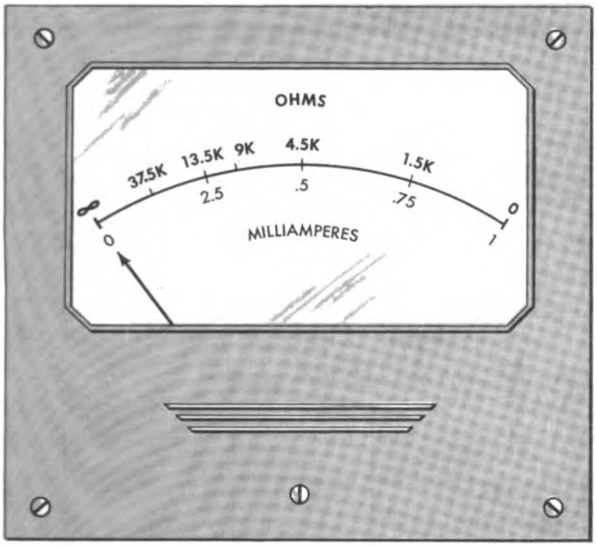


Circuit of an Ohmmeter

Most ohmmeters have more than one scale. Additional scales are made possible by using various values of limiting resistors and battery voltages. Some ohmmeters have a special scale called a *low ohm* scale for reading low resistances. A *shunt-type ohmmeter* circuit is used for this scale.

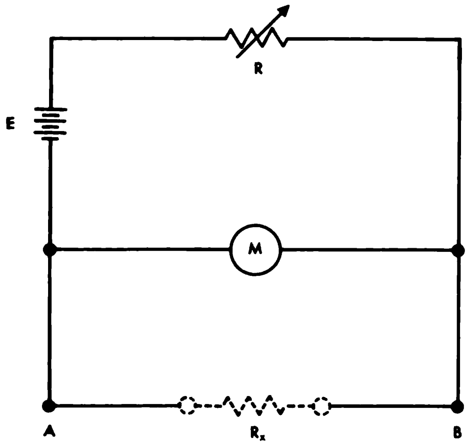
Shunt-Type Ohmmeter

Shunt-type ohmmeters are used to measure small values of resistance. In the schematic circuit shown on page 77, E is applied across a



A Typical Ohmmeter Scale

limiting resistor R and a meter movement in series. Resistance and battery values are chosen so that the meter movement deflects full scale when terminals A and B are open. When the terminals are short-circuited, the meter reads zero; the short circuit conducts all the current around the meter. The unknown resistance R_X is placed between terminals A and B in parallel with the meter movement. The smaller the resistance value being measured, the less is the current through the meter movement; the greater the value of the resistance, the more the current through the meter movement. The scale of a shunt-type ohmmeter is therefore a *direct* scale; that is, values increase from left to right.



Shunt-type Ohmmeter

The value of the limiting resistor R is usually made large compared to the resistance of the meter movement. This keeps the current drawn from the battery practically constant. Thus the value of R_X determines how much of this constant current flows through the meter and how much through R_X.

Note that in a shunt-type ohmmeter current is always flowing from the battery through the meter movement and the limiting resistor. Therefore, when using an ohmmeter with a low-ohm scale, *do not leave the switch in low-ohm position.*

VOLT-OHM MILLIAMMETER

Each of the meters discussed consists of a basic meter movement combined with additional devices necessary for its particular purpose. Shunts are used for ammeters, multipliers for voltmeters, and resistors and batteries for ohmmeters. A voltmeter, ohmmeter, and ammeter may be combined in one unit using only one meter movement. Switches, jacks, and the necessary additional devices are included to make such a combination volt-ohm-milliammeter or multimeter. By proper arrangement of parts, the multimeter can be built into a small, compact unit.

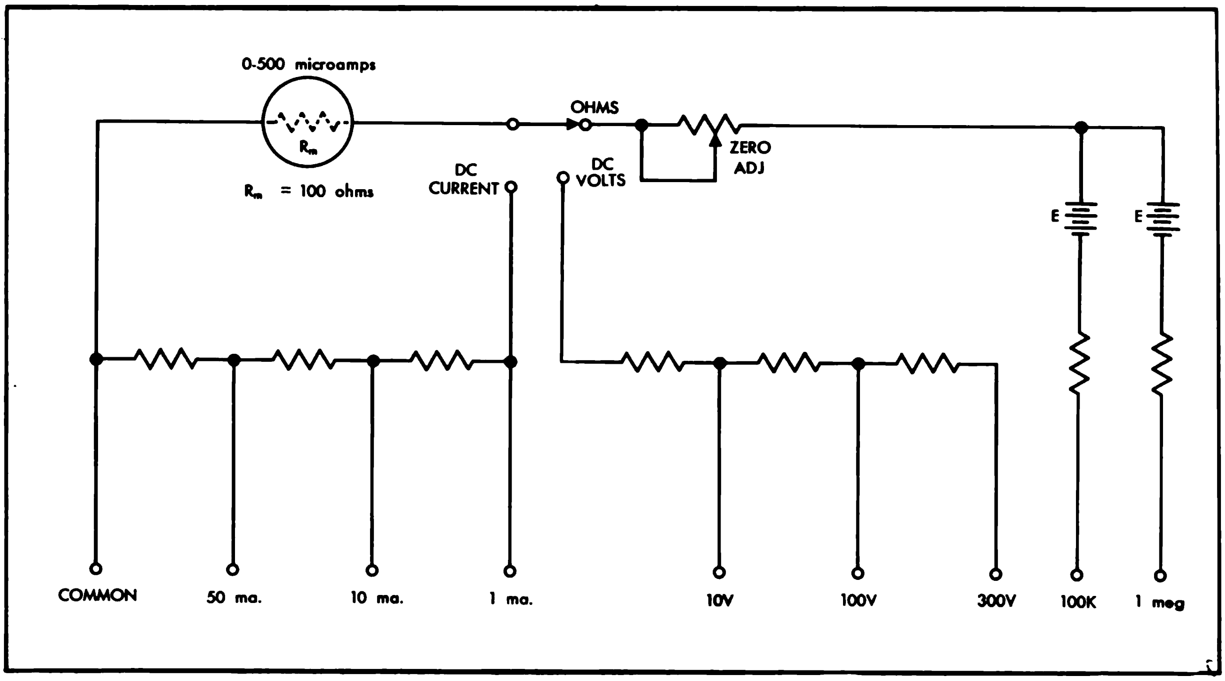
A typical multimeter circuit is shown in the illustration on page 78. A three-range milliammeter, a three-range voltmeter with a sensitivity of 2000 ohms-per-volt, and a two-range ohmmeter are combined in this circuit. A 0-500 microampere meter movement with a resistance of 100 ohms is the basic meter movement for this multimeter.

MEASURING AC VOLTAGE

Since the D'Arsonval movement can measure only direct current, a device which changes AC to DC must be used with this type meter movement if you wish to measure AC. Such a device is called a rectifier. There are many types of rectifiers, some of which are discussed in chapter 19.

Using a Rectifier for Measuring AC

A copper-oxide rectifier is generally used with DC meter movements to measure AC. This allows current to flow through the meter in only one direction. As shown in the illustra-



A Typical Multimeter Circuit

tion below, the copper-oxide rectifier consists of copper-oxide discs separated alternately by copper discs and fastened together as a single unit. Current flows more readily from copper to copper oxide than from copper oxide to copper. When AC is applied, therefore, current flows in only one direction, yielding a pulsating DC output as shown by the output waveshapes in the illustration below. This current can then be measured as it flows through the meter movement.

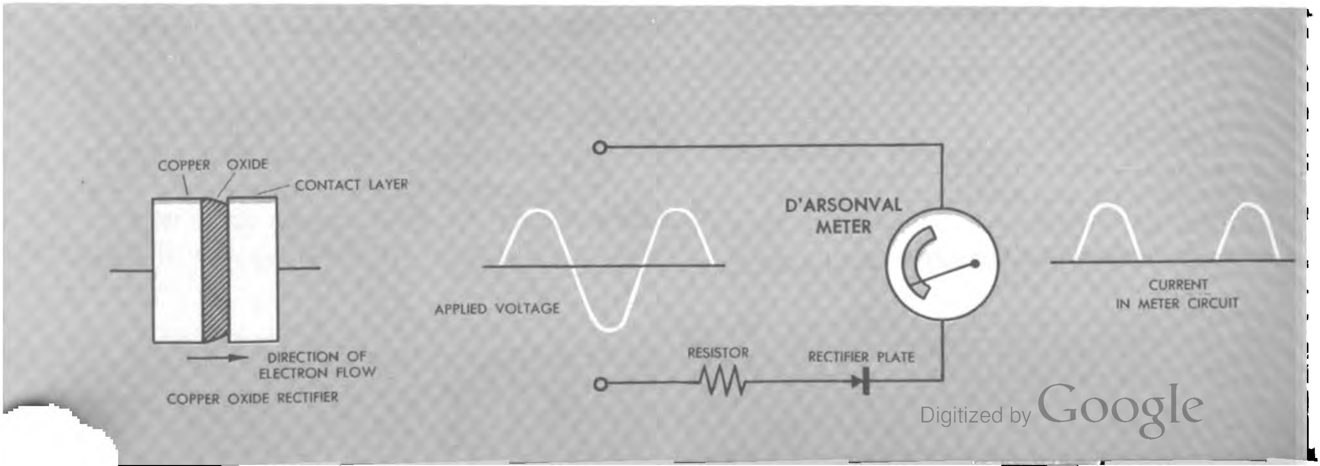
In some AC meters selenium or vacuum tube rectifiers are used in place of the copper-oxide rectifier. The principle of operation, however, is the same in all meters employing rectifiers.

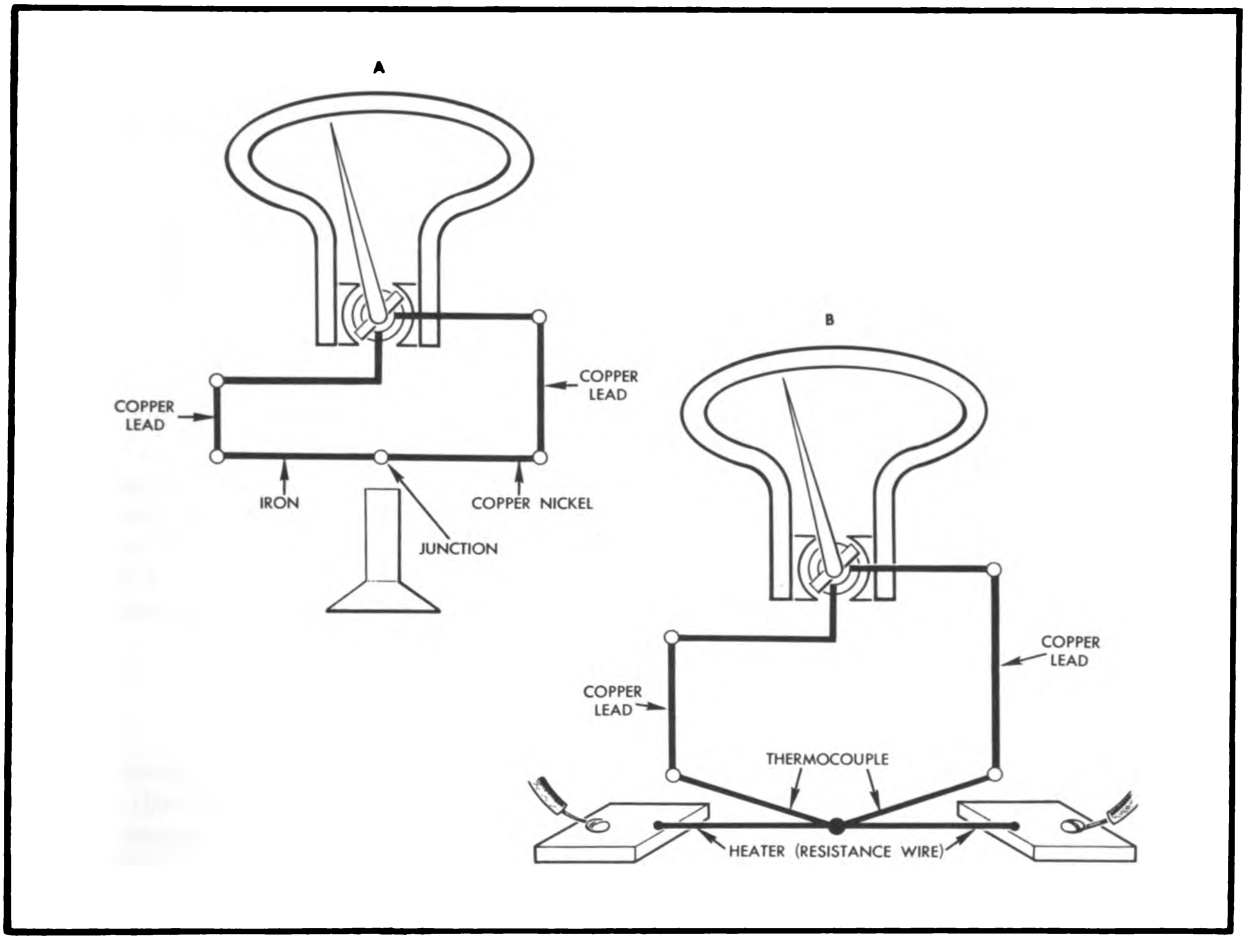
Using a Thermocouple for Measuring AC

The D'Arsonval movement can also be used with a thermocouple to measure AC or DC. A *thermocouple* is a combination of two dissimilar metals connected at one end and open at the other. If the point of contact is heated as shown at A in the illustration on page 79. a potential difference exists between the open ends. When a conductor is connected between the open ends, current flows. The direction of current flow is determined by the combination of metals used.

In meter applications, a D'Arsonval movement is used as the conductor between the open ends of the thermocouple and the flame is

Half-wave Meter Circuit





Thermocouple

replaced by a resistance wire heater as shown at B. As current flows through the resistance wire, the heat developed is transferred to the contact point and develops an EMF which causes current to flow through the movement. The coil rotates and causes the pointer to move over a calibrated scale. The amount of coil movement is dependent upon heat which varies as the square of the current. Therefore, a square-law scale results. In some meters, however, the scale is made more linear by placing the permanent magnet pole pieces off center. Thermocouple meters find extensive use in AC measurements.

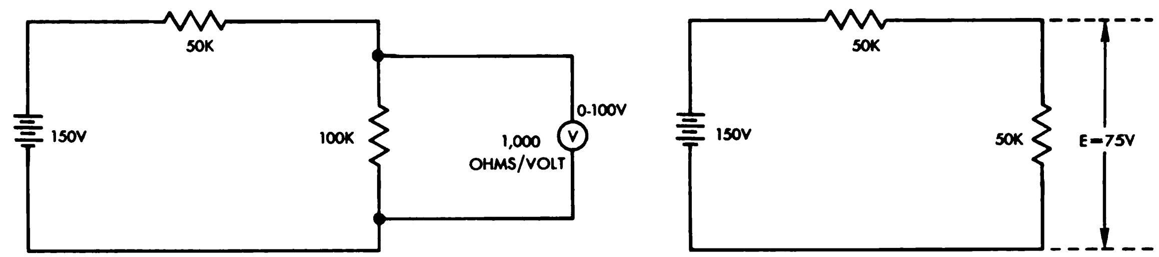
METERS AFFECT CIRCUITS

Because meters have resistance, they affect the circuits in which they are placed. The sensitivity of a meter is an indication of how much the meter affects the circuit and, therefore, the accuracy of the reading. When you use a meter, keep the following in mind:

An ammeter must have *considerably less resistance* than the circuit with which it is in series. A voltmeter must have *considerably more resistance* than the circuit with which it is in parallel.

The fact that an ammeter must have very low resistance for accurate reading is obvious. Since the meter is placed in series, it increases the total resistance of the circuit and causes less current to flow. This results in a reading which is low. The greater the resistance of the meter, the lower is the reading.

The effect of placing a voltmeter in a circuit may be better understood by examining the illustration on page 80, Loading Circuit with Voltmeter. Before the meter is connected, the voltage



Loading Circuit with Voltmeter

across R_2 is 100 volts. If the meter is connected in parallel on the 0-100 volts scale, you are effectively placing another 100 K resistor in parallel. (The resistance of the meter on any scale is obtained by multiplying its sensitivity by the full-scale voltage reading.) The total resistance across R_2 is now 50 K and the voltage is 75 volts. The meter has loaded the circuit and caused an inaccurate reading.

If the same meter is used on the 500-volt range, the resistance placed in parallel with R_2 is 500 K. The total resistance across R_2 becomes approximately 83 K and the voltage approximately 91 volts. If a 20,000 ohms-per-volt meter is used, the circuit is not loaded as much and the reading is more accurate.

Chapter 12

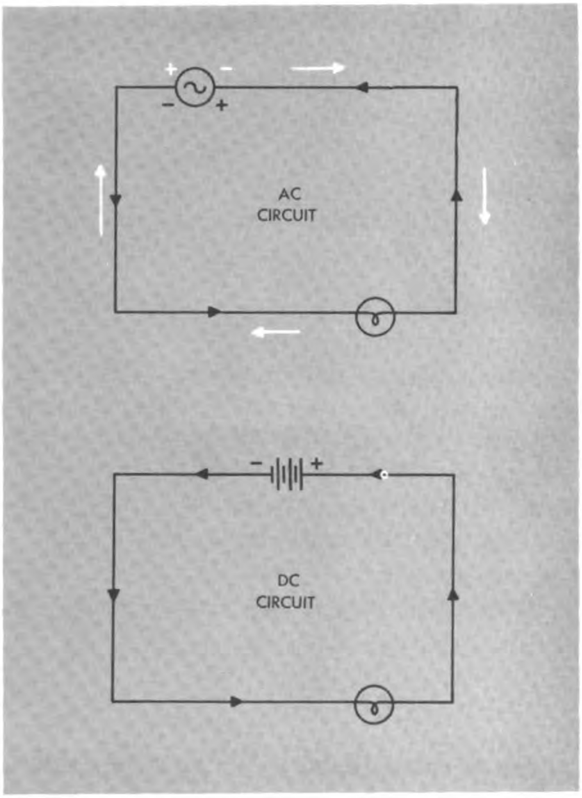
Basic Concepts of Alternating Current

Approximately 90 percent of all electric power that is consumed throughout the world is developed by alternating-current generators. The principal reason for this is that less power is wasted in the transmission of alternating current than in the transmission of direct current. An alternating voltage can be *stepped up*, with a consequent decrease in the amount of alternating current, for transmission over a long power line. At the consumer end of the line the voltage can be *stepped down* to a suitable value with a consequent increase in current. Since the amount of current is decreased for transmission, the amount of power loss in the line is also decreased. This over-all process makes it practical to locate the power plant near a source of potential energy such as a waterfall or a coal field and to deliver the electric power to a suitable manufacturing site or a distant city by means of electric power lines.

There are several reasons why alternating current also plays an important part in communications. *First of all*, whatever means of communications you are using, you are dealing with variations in the transmitting medium. For example, voice is carried by sound waves, and sound waves are variations in air pressure. To make it possible to communicate by telephone or radio, sound waves must be made to cause electrical variations. Alternating currents are electrical variations. Therefore, you can say that the intelligence in the sound waves is transmitted by telephone or radio as a type of alternating current. *Secondly*, electric energy radiates from a circuit in which there is an alternating current flowing. This makes it possible to transmit intelligence in the form of voice or code messages, and still pictures, without the use of connecting wires between the transmitting and receiving stations.

AC AND DC COMPARED

A direct current consists of a number of electrons moving from a point of low potential to a point of high potential. For example, when the circuit of a flashlight is completed by the switch, electrons continue to move from the negative terminal toward the positive terminal through the filament of the bulb, as long as the switch is ON. In an alternating-current circuit, the electrons move through the circuit in one direction for a short period of time and then they move back in the other direction for a like period of time, changing direction at definite intervals. In this case, the electrons also move toward a point of higher potential, but this



AC and DC Circuits

point changes from one side of the circuit to the other. In other words, the polarity of the applied alternating voltage changes, causing the direction of current flow to change.

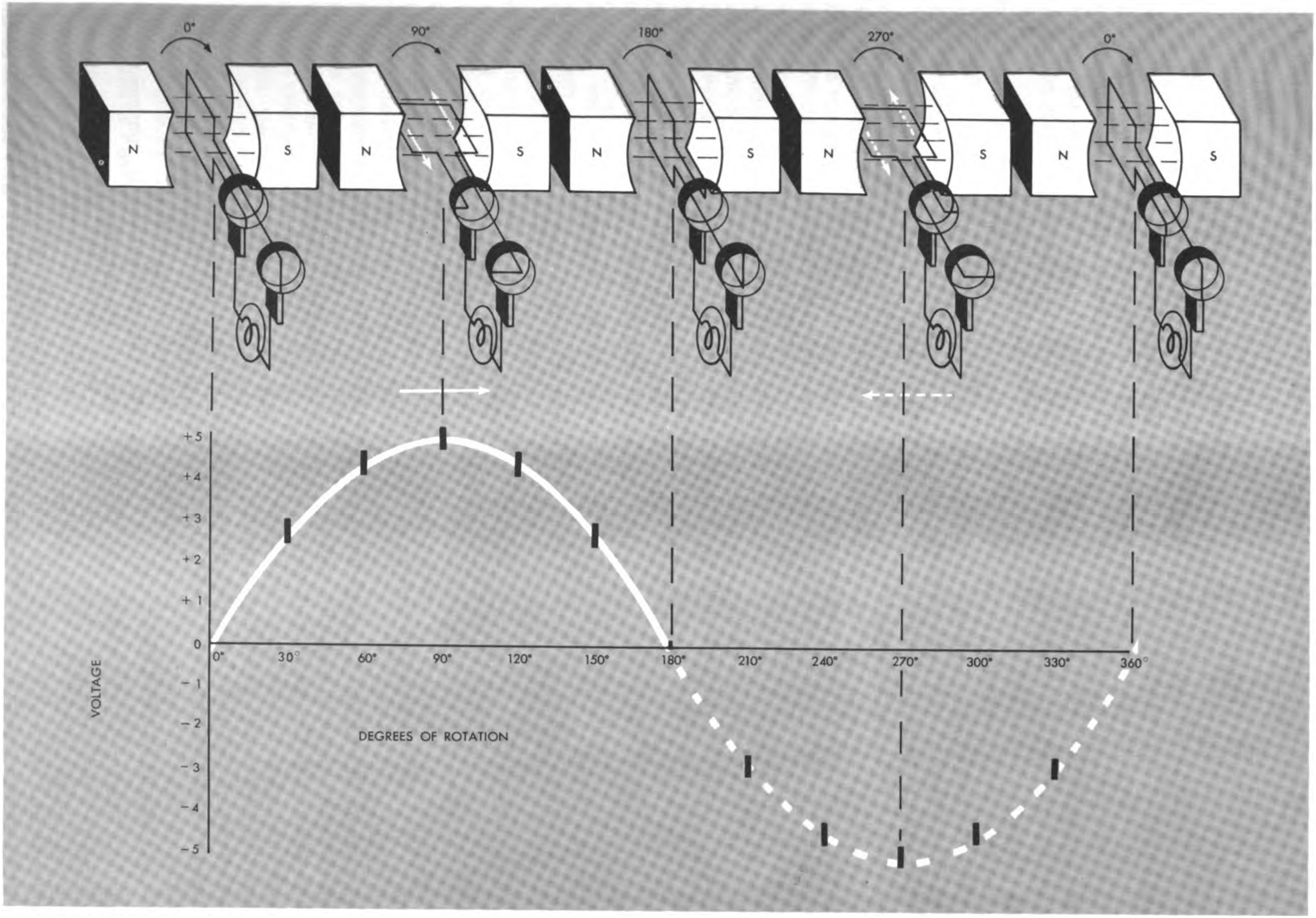
DEFINITION OF AC

An alternating current is a current that is continuously varying in magnitude and periodically changing direction. The rate at which the current changes direction of flow is twice the frequency of the alternating current. If 120 reversals of current occur each second, the *frequency* of the alternating current is $\frac{120}{2}$ or 60 cycles per second. (The term *cycle* means one complete set of events or phenomena that occur periodically.) The standard frequency for the generation and transmission of electric power in the United States is 60 cycles per second. This is abbreviated 60 cps. You have, no doubt, seen the following sentence on the nameplate of an electric appliance: "Use on 60-cycle AC only." This means that the appliance was designed to operate on this specific frequency of alternating current. Operation on some other *power frequency* or on direct current would probably damage the appliance. A power frequency of 25 cps is used in some localities, and in various foreign countries, the power frequency may be almost any frequency from 25 cps to 125 cps.

Frequencies higher than power frequencies are also used in electronic equipment. Frequencies from 16 to 16,000 cps are called *audio frequencies* since these frequencies, when reproduced as sound, can be heard by the human ear. Frequencies above 16,000 cps are called *radio frequencies* because they can be efficiently radiated into space. Different ranges of radio frequencies are more adaptable to some applications than to others. For example, frequencies from 20,000 to 100,000 cps (20 kilocycles to 100 kilocycles) are most useful for long distance communications or transoceanic communications. Extremely high frequencies such as those in the band from 10,000 to 30,000 megacycles per second are most applicable to radar systems. One kilocycle per second equals one thousand cycles per second, and one megacycle per second equals one million cycles per second. Kilocycle per second is abbreviated kc and megacycle per second is abbreviated mc. Frequencies over 30,000 mc are being used for various types of experimental work.

GENERATION OF AN ALTERNATING VOLTAGE

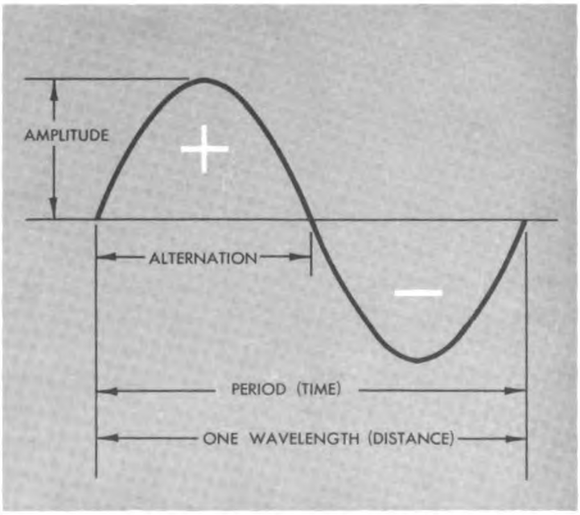
A simple AC generator, called an *alternator*, consists of a single-turn coil of wire, a permanent magnet, two slip rings, two brushes, and a brush holder. A *slip ring* is a continuous metallic ring, unlike the segmented commutator of a DC generator. Each end of the coil of wire is connected to a slip ring and the two slip rings are insulated from each other. The coil of wire and the slip rings make up the armature or *rotor* and this rotor is mounted between the poles of the magnet so that it is free to rotate on its axis. The slip rings are in a position such that their axes are in line with the axis of the coil of wire. External connections to the slip rings are made by means of the brushes that are held stationary by the insulating brush holder and make continuous contact with the slip rings.



Generation of AC Voltage

The illustration on page 83 pictures a simple AC generator with its rotor in various positions. The sine wave of voltage generated in one complete revolution is also shown. An explanation of how the rotating coil generates this type of voltage was presented on page 56 of this manual. Before proceeding, reread and study this explanation. Be sure you understand how an AC voltage is generated and why a sine wave is an effective representation of this type of voltage.

The frequency of the output voltage of the simple alternator is determined by the *speed of rotation* of the rotor. For example, if the rotor makes 60 complete rotations per second, the output frequency would be 60 cps.



AC Terminology

The amount of voltage generated by the simple alternator depends primarily upon the *strength of the field.* The speed of cutting also has a bearing upon the amount of voltage as well as on the frequency. These factors are the same as those that affect the amount of voltage generated by a DC generator.

AC TERMINOLOGY

You have already become familiar with the terms *frequency* and *cycle* with respect to an alternating voltage. There are a number of other terms that are used in conjunction with alternating currents and voltages. Since the study of nearly all electronic equipment involves the theory of alternating current, you should become familiar with these additional terms illustrated on this page.

Alternation

The positive half of an alternating voltage is called an *alternation.* Likewise, the negative half is also called an alternation. It is sometimes convenient to refer to these as the positive alternation and negative alternation respectively. Since there is a positive and a negative alternation for each cycle, the number of alternations per second is equal to the frequency multiplied by two.

Period

The time, in seconds or fractions of a second, required to generate one cycle of alternating voltage is called the *period.* There is a definite relationship that exists between the frequency and the period of an alternating voltage. In the case of a 60-cycle voltage, 60 complete cycles are generated each second. Therefore, it requires 1/60th of a second to generate one cycle. This short interval of time is called the period of the 60-cycle voltage. The relationship between period and frequency can be conveniently expressed by the formula $t=\frac{1}{f}$ where t is the period in seconds and f is the frequency in cycles per second.

Wavelength

The speed of an *individual electron* through a conductor depends upon a number of factors such as the type of material of which the conductor is made and the potential difference across the conductor. However, the speed of an *electric current* through a conductor and likewise the speed of radiation of a radio wave from an antenna is approximately *186,000 miles per second* or *300 million meters per second.* (A meter is equal to 39.37 inches.) If the speed or velocity of the electric current or the radio wave is divided by f, its frequency, the quotient is the physical length of one wave of the *wavelength.* This relationship is conveniently expressed by the formula,

$$\lambda=\frac{v}{f}$$

where λ (Greek letter lambda) is the symbol for wavelength, v is the velocity, and f is the frequency. A wavelength is defined as the distance between two successive like points on the wave. For example, the distance between two successive points at which the alternating voltage is passing through zero in a positive direction is one wavelength. The distance between two successive points where the wave is ½ maximum positive, but decreasing, is also one wavelength, and so on.

Substituting the power frequency of 60 cycles per second in the above formula, you will find the wavelength to be 5 million meters or 3,100 miles. Thus, if a hypothetical power line is many thousand miles long, points at which the voltage is maximum positive, for example, are 3,100 miles apart. The wavelengths of radio frequencies are much shorter than those of power frequencies. When a radio frequency of 100 million cycles per second or 100 megacycles is substituted in the formula, the wavelength is found to be 3 meters or approximately 9.84 feet. If you want to find the wavelength in meters, express the velocity in meters per second (300 million.) If you desire to find the wavelength in feet, express the velocity in feet per second. To do this multiply the velocity of 186,000 miles per second by 5,280, since there are 5,280 feet in a mile.

Amplitude

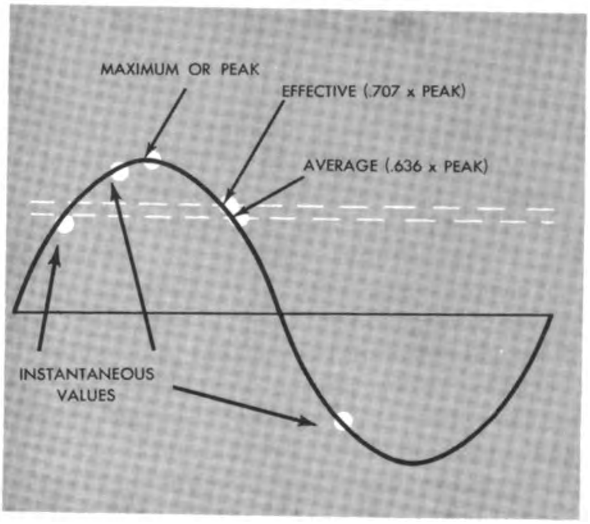
The *amplitude* of a wave is the distance from zero to the maximum value. The positive amplitude and the negative amplitude are ordinarily the same. In the graph of the typical simple alternator output voltage shown on page 83, the amplitude is 5 volts.

ALTERNATING-CURRENT VALUES

In discussing alternating current and voltage, it is often necessary to express the current and voltage in terms of maximum or peak values, effective values, average values, or instantaneous values. Each has a different meaning and is used to denote a different amount of current or voltage as shown in the illustration, AC Values.

Maximum or Peak

The term maximum has already been used in the discussion of the simple alternator. The *maximum value* of an alternating voltage was indicated as being the highest voltage that was reached on the positive or negative alternation. A similar term that is used to express this value is peak. The *peak value* is the highest value of voltage or current reached during either the positive or negative alternation.



AC Values

Instantaneous

The *instantaneous value* of voltage or current is the value of voltage or current at one particular instant. This value may be zero if the selected instant is the time during which the polarity of the voltage is changing. It may also be the same as the peak value if the selected instant is the time that the voltage or current stops increasing and starts decreasing. There may be any number of instantaneous values between zero and the peak value.

Average

The *average value* is the average of all the instantaneous values during one alternation. Since the voltage increases from zero to peak and decreases back to zero during one alternation, the average value must be some value between those two limits. Mathematically, *the average value is equal to .636 times the peak value* ($E_{av} = .636\ E_{peak}$). Do not confuse this definition of average value with the average over a complete cycle. Since the voltage goes positive during one alternation and negative during the next alternation, the average of a complete cycle is zero.

Effective

The *effective value* is the value of alternating current or voltage that will cause the same amount of heat to be produced in a resistive circuit that is caused by a DC voltage or current of the same value. *The effective value is equal to .707 times the peak value* ($E_{eff} = .707\ E_{peak}$). The effective value is also called the root mean square (nm) value. Mathematically, this value is equal to the square root of the sum of the squares of all the instantaneous values of current or voltage during one alternation. The voltage rating of an alternator or the voltage available in an AC outlet in your home or in the laboratory work bench is expressed in terms of effective value. In other words, 110 volts is the effective value of the voltage available from a "110-volt line."

Converting Values

In studying and working with alternating current you may have to convert a given or measured value of voltage or current to another value, for example, you may have to convert an effective value to its corresponding peak value.

Solving the effective voltage formula for peak voltage,

$$E_{eff} = .707\ E_{peak}$$

$$E_{peak} = \frac{E_{eff}}{.707}$$

$$E_{peak} = 1.414\ E_{eff}$$

Substituting 110 volts for the effective voltage, you can find the peak voltage of a "110-volt line" as follows:

$$\begin{aligned} E_{peak} &= 1.414\ E_{eff} \\ &= 1.414 \times 110 \\ &= 155.54 \text{ volts} \end{aligned}$$

In other words, the *amplitude* of the sine-wave voltage of a "110-volt line" is 155.54 volts. The *average value* of the sine-wave voltage of a "110-volt line" is equal to .636 times this peak value or 98.92 volts.

If you desire to convert an average value to a peak value, solve the average-voltage formula for peak voltage.

$$E_{av} = .636\ E_{peak}$$

$$E_{peak} = \frac{E_{av}}{.636}$$

$$E_{peak} = 1.572\ E_{av}$$

Now suppose you wish to convert an average value to an effective value.

$$.636\ E_{peak} = E_{av}$$

$$.707\ E_{peak} = E_{eff}$$

Since the average value is less than the effective value, E_{av} times some number greater than one must equal E_{eff}. Call this unknown number Y. Then

$$E_{av} \times Y = E_{eff}.$$

Substituting the above values for E_{av} and E_{eff},

$$.636\ E_{peak} \times Y = .707\ E_{peak}$$

$$Y = \frac{.707\ E_{peak}}{.636\ E_{peak}}$$

$$Y = 1.11$$

Substituting this value of Y above,

$$1.11\ E_{av} = E_{eff}$$

If it is desired to convert an effective value to an average value,

$$1.11\ E_{av} = E_{eff}$$

$$E_{av} = E_{eff} \times \frac{1}{1.11}$$

$$E_{av} = .9\ E_{eff}$$

Note that all of the above conversions are made from the average-voltage and effective-voltage formulas. Once you have learned these two, it merely requires simple algebraic manipulation to find the other formulas.

Chapter 13

Inductors

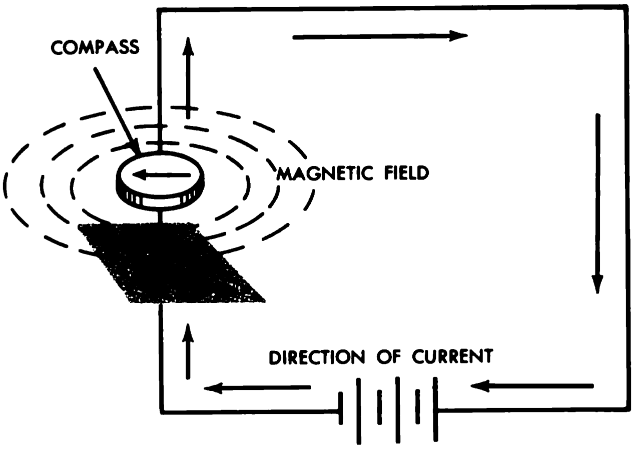
In direct-current circuits, the opposition to the flow of current is called *resistance*. Conductors, such as copper wires, have small amounts of resistance, but most of the resistance in a DC circuit is the resistance of a circuit component such as a light bulb, an electric motor, a relay solenoid, or a wire-wound or carbon resistor. The unit of resistance is the ohm. A resistance of one ohm limits the current flow to one ampere when the applied potential is one volt.

In alternating-current circuits, the opposition to the flow of current is called *impedance*. The unit of impedance is the ohm. An impedance of one ohm limits the flow of alternating current to one ampere when the applied AC potential is one volt. If a resistor is the only circuit component connected across a source of AC voltage, the impedance and the resistance of the circuit are the same. The impedance in an AC circuit may also be due to a component called a *reactor*. There are two distinct types of reactors . . . *inductors* and *capacitors*. These components are called reactors because they do not dissipate electric energy in the form of heat, as do resistors, but alternately store energy and then deliver this energy back to the circuit. Complex AC circuits may consist of various combinations of resistors and reactors.

This chapter deals with inductors, their physical and electrical characteristics, and their effects in AC circuits. Typical components and practical methods of checking inductors are also described.

AC circuits consisting of various combinations of resistors and/or reactors are discussed in detail in chapters 16 and 17. The simplest AC circuit is the series circuit, which is made up of reactors and/or resistors connected in such a way that the total current from the power source flows through each component. In a parallel circuit, the total current from the power source divides, some flowing through each branch. Series-parallel circuits are combinations of series and parallel circuits.

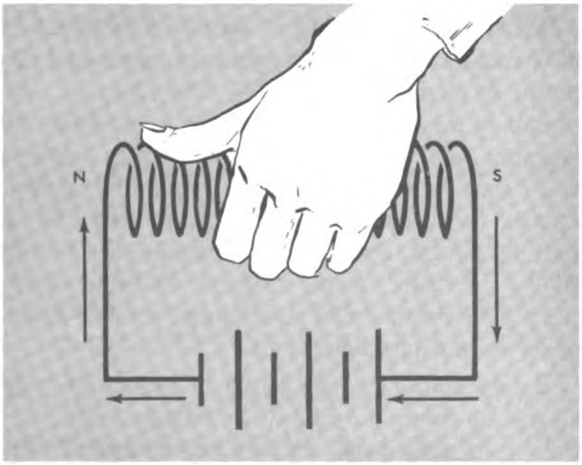
Before beginning the study of inductors, you should be thoroughly familiar with some electromagnetic principles, studied previously. When current flows through a conductor, a magnetic field is established around the conductor. If you



Magnetic Field Around a Conductor

wrap the fingers of your left hand around a current-carrying conductor with your thumb pointing in the direction of current ~~flow,~~ your fingers will point in the direction of the magnetic field around the conductor. If you hold a small compass near the conductor, the compass needle will be aligned with this magnetic field.

When current flows through a coil of wire, the magnetic fields around the individual turns combine so that one end of the coil becomes a strong north magnetic pole and the other becomes a strong south magnetic pole. When you wrap the fingers of your left hand around the coil in the direction of current flow, your extended thumb points in the direction of the north pole as shown in the illustration below.



Magnetic Field Around a Coil

As current *increases* through the coil, the magnetic flux lines around each individual turn expand, and pass or cut across the adjacent turns. Using the left-hand rule, you will find that the effect of these individual expanding fields is to tend to cause an induced current to flow in the opposite direction to that of the original current that caused the expanding field. This phenomenon is stated by an abbreviated form of Lenz's law as follows: *When a voltage is induced in a coil as a result of any variation of the magnetic field with respect to the coil, the induced voltage is in such a direction as to oppose the current change that caused the magnetic variation.* From this, you can also say that a *decrease* in the current flow through a coil also causes an induced voltage that opposes the decrease in current. This opposition to a change in current, illustrated on page 89, is called *back EMF* or *counter EMF* and gives rise to the electrical property called inductance.

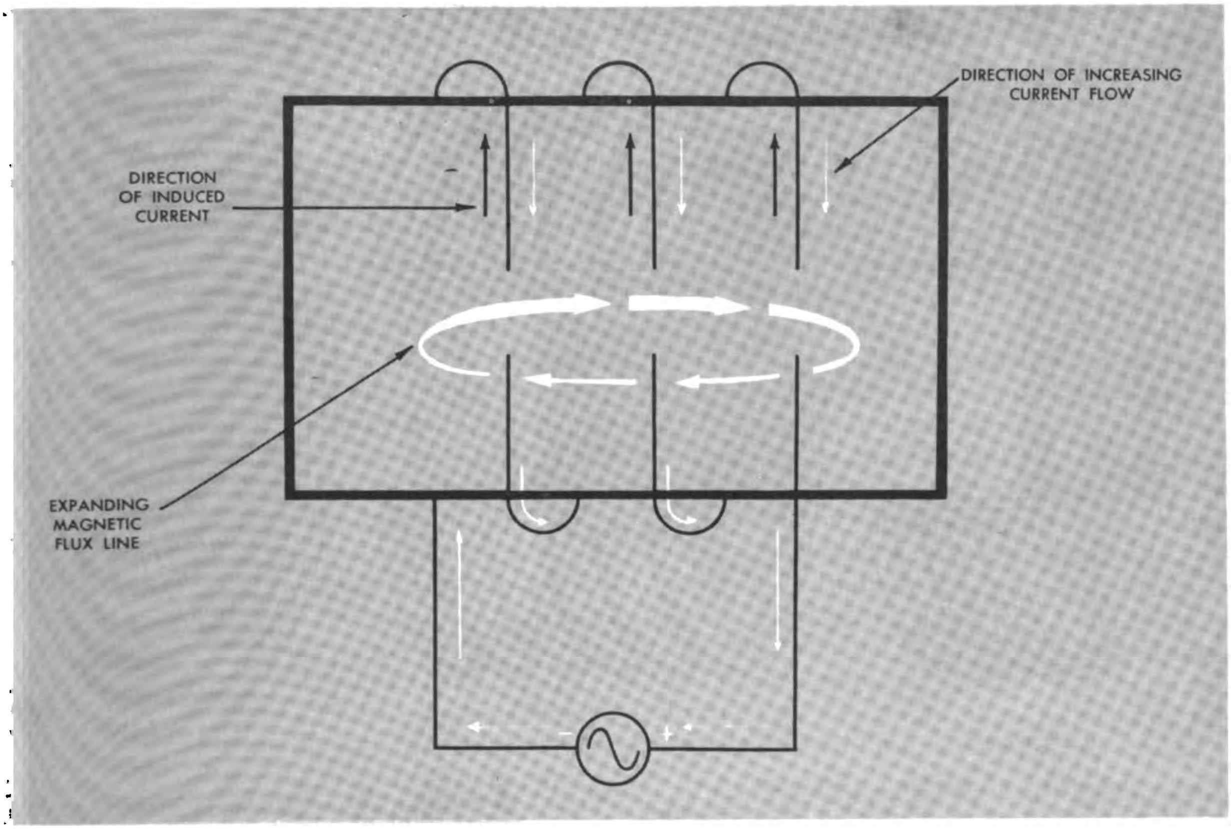
Inductance is defined as that property of a circuit which opposes any change in the rate of current flow. This property has little effect in a direct-current circuit, since there is no variation in current and consequently no change in the magnetic field to induce a back EMF. When an alternating current flows through a coil, however, the individual fields around the turns are constantly expanding and collapsing, producing a continuous back EMF. The symbol for inductance is L, and the basic unit is the *henry.* *A coil is said to have an inductance of one henry when a change in current of one ampere per second causes an induced back emf of one volt.*

FACTORS AFFECTING INDUCTANCE

Any conductor that is carrying an alternating current has a small amount of inductance. However, an inductor is usually a coil or wire. The factors that affect the amount of inductance of a coil of wire are the *number of turns,* the *type of core,* the *spacing and method used in winding the coil,* the *diameter of the coil,* and the *ratio of the diameter to the coil length.*

Number of Turns

Assume that a coil has a given number of turns and is carrying a definite amount of alternating current. If some turns are added to the original turns, the current flow through the


DIRECTION OF INCREASING CURRENT FLOW
DIRECTION OF INDUCED CURRENT
EXPANDING MAGNETIC FLUX LINE

Back EMF

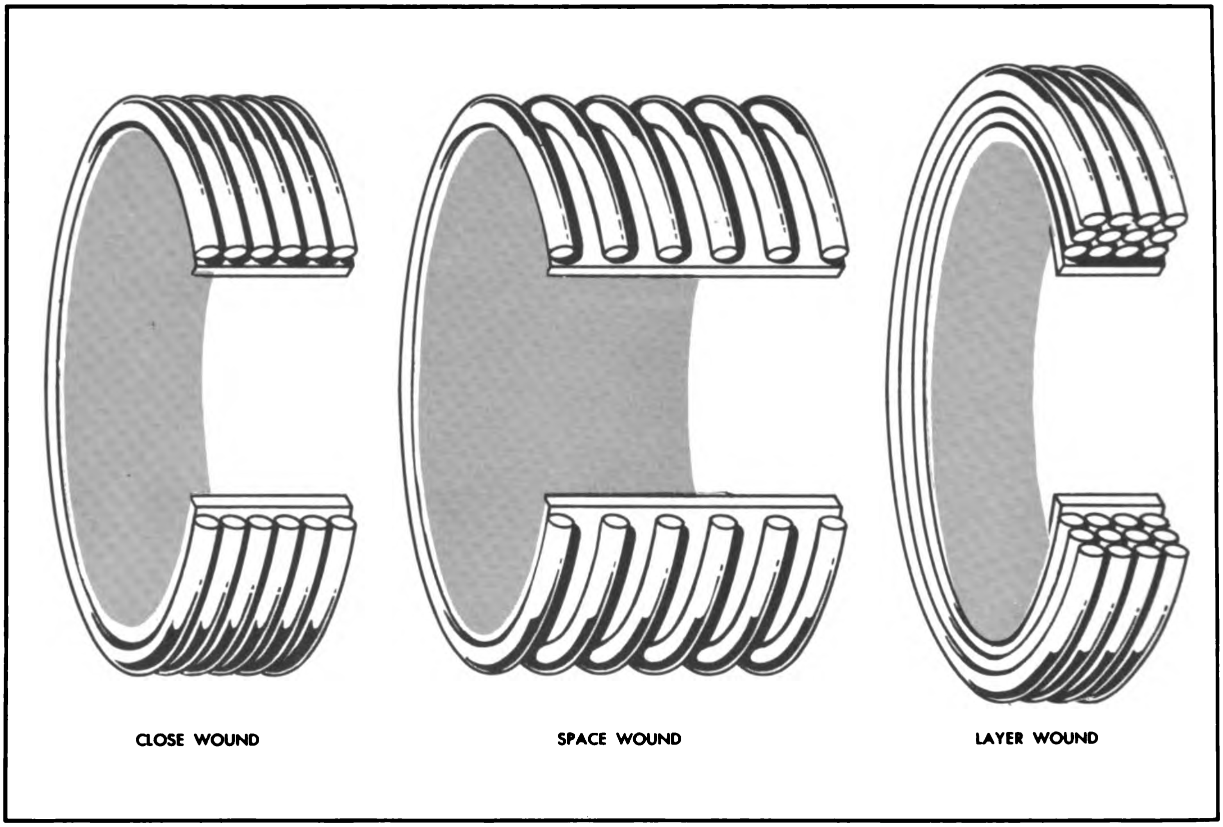
total number of turns will create a *stronger magnetic field.* This stronger field will induce a greater back EMF in the original turns. In addition the stronger field also cuts the *added turns.* Both of these factors cause an increase in inductance. Since the effects of these two factors combine, you can say that the inductance of a coil is directly proportional to the *square* of the number of turns. For example, if the number of turns is doubled (multiplied by 2), the inductance will be 2^2 or 4 times the original inductance.

Type of Core

The inductance of a coil is directly proportional to the permeability of the core. (Permeability is the ease with which a material passes magnetic lines of force.) Coils having nonmagnetic cores have much less inductance than coils with magnetic cores, assuming that the number of turns and other factors are equal. Typical nonmagnetic-core coils are those that are wound on hollow porcelain cylinders and hollow cardboard cylinders. The core material in these coils is air. Typical magnetic-core coils are those that are wound on cylindrical or rectangular pieces of soft iron (iron-core coils). Since an iron core has much less opposition to the lines of force than an air core, the magnetic fields around the individual turns of the iron core coil are stronger than those around the individual turns of an air-core coil. This causes a greater induced back EMF in the adjacent turns of the iron-core coil, resulting in an over-all greater inductance.

Spacing and Method of Winding

There are a number of methods used in winding coils. Some coils are wound so that adjacent turns touch each other. These are called *close-wound coils.* Other coils are wound so that there is a certain amount of space between the adjacent windings. These are called *space-wound coils.* They have less inductance than the close-wound coils, since the magnetic fields of the individual turns must travel farther to cut the adjacent turns. The third common type is the *layer-wound coil.* This type has more inductance than either



Types of Coil Windings

of the previously mentioned types since the individual magnetic fields are more concentrated with respect to the adjacent turns. Each type has special applications which you will learn about in the study of communications and radar equipment.

Diameter of the Coil

The greater the diameter of a coil the greater its inductance. Physically, it requires more wire to construct a coil of large diameter than one of small diameter with an equal number of turns. Consequently, more lines of force exist to induce a back EMF in the coil with the larger diameter.

Ratio of Diameter to Coil Length

The greater the ratio of the coil diameter to the coil length, the greater the inductance. If the diameter is relatively large and the coil length relatively short, the field around a single winding will induce a certain amount of back EMF in a number of other windings and perhaps in all the other windings. On the other hand, if the ratio is small, the field around an individual winding will induce a back EMF in fewer of the total windings since they are distributed over a greater distance (length). Coil winding tables, available in certain radio handbooks, give the inductances of coils having various diameter-to-length ratios.

SELF AND MUTUAL INDUCTANCE

When the field of an inductor induces a back EMF in the inductor itself, it is said that the inductor possesses inductance or, more specifically, *self-inductance.* The voltage of self-induction depends upon the inductance of the coil and the rate at which the current changes through the coil. This voltage is opposite in polarity to the applied voltage and is represented mathematically by the formula,

$$e_i = -L\frac{di}{dt}$$

where e_i is the induced back EMF at a given instant, L is the inductance of the coil, and

$\frac{di}{dt}$ represents the change in current (i) per unit of time (t). The induced back EMF is greatest when the amount of current change per unit time is the greatest or when the current is changing the fastest. The minus sign signifies that the back EMF opposes the applied voltage.

When two inductors, or coils, are placed close enough together so that the magnetic field of one coil cuts the turns of the other, it is said that *mutual inductance* exists between the coils. In common electronics terminology, the coils are said to be coupled. If two coils (L_A and L_B) are connected in series and are physically close enough together so that coupling exists between them, the field from coil L_A affects coil L_B and likewise the field from coil L_B affects coil L_A. If the individual inductances of the two coils are equal, the effect of one upon the other is equal since they are connected in series and carry the same amount of current. The amount of inductance in coil L_B due to the field from coil L_A is designated M, and the amount of inductance in coil L_A due to the field from coil B is also M since their effects are equal. The total inductance of the two coils is therefore equal to the sum of each individual inductance and the effects of each on the other or

$$L_T = L_A + L_B \pm 2M$$

PHASE RELATIONSHIP IN AN INDUCTOR

An inductor has a certain amount of resistance since it is made up of a number of turns of wire. However, the amount of resistance in many types of inductors is so small that it can be neglected. When the resistance is neglected, you can say that the inductor has *pure inductance.*

When you connect a resistor to a source of alternating voltage, the current through the resistor remains in phase with the voltage across the resistor. This means that the current is zero when the voltage is zero, and the current is maximum when the voltage is maximum. However, when you connect an inductor to a source of alternating voltage, to form an inductive circuit, the current through the inductor lags the voltage across the inductor. This means that the current does not reach its maximum value until after the voltage has reached its maximum value. *The current lags the voltage in a purely inductive circuit by 90°.*

Suppose that current is starting to flow through a coil. The current causes an expanding magnetic field which causes a back EMF to be induced in the coil. The polarity of the back EMF is opposite to the polarity of the applied voltage across the coil, and it therefore tends to oppose an increase in current. The result is that the rise in current flow is caused to lag behind the rise in voltage. As current increases through the coil, energy is stored in the magnetic field. When the applied voltage starts to decrease from maximum positive, this energy is returned to the circuit in the form of current flow. This current flow is actually greater than the current flow due to the applied voltage. When the circuit is purely inductive, maximum current flows at the instant that the voltage has decreased to zero from its positive direction. A similar action occurs on the negative portion of the applied voltage, and maximum current flows in a purely inductive circuit in the reverse direction at the instant that the voltage has decreased to zero from its negative direction.

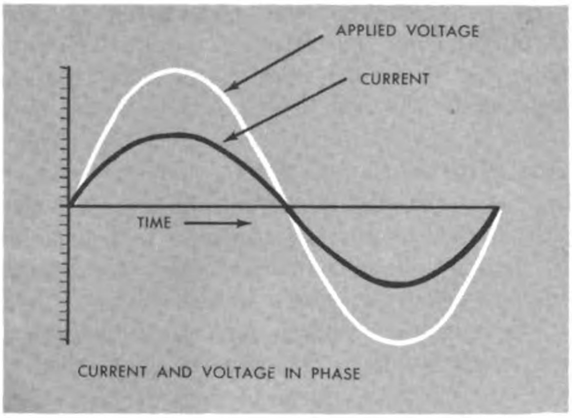
The phase relationships in a resistive circuit and a purely inductive circuit are shown in the illustrations on page 92. Note that the current is zero when the voltage is zero and the current is maximum when the voltage is maximum in the resistive circuit. In the inductive circuit, however, the current reaches zero 90° after the voltage reaches zero. The back EMF always opposes the applied voltage. In a purely inductive circuit, it is maximum when the current is zero since the rate of change of current and speed of cutting of the magnetic field are the fastest at this time.

INDUCTIVE REACTANCE

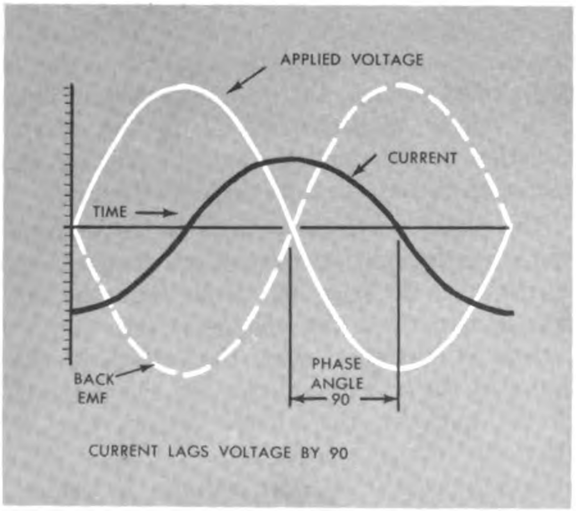
The opposition of an inductor or a coil to the flow of alternating current is called *inductive reactance.* The amount of inductive reactance of a coil depends upon the inductance of the coil, the frequency of the alternating current, and a constant value of 2π which is equal to 6.28 (approximately). Stated as a formula,

$$X_L = 2\pi f L$$

where X_L is the inductive reactance in ohms, f is the frequency of the applied alternating current in cycles per second, L is the inductance of the coil in henries, and the Greek letter π equals 3.1416.



Waveforms for Resistive Circuit



Waveforms for Purely Inductive Circuit

Example

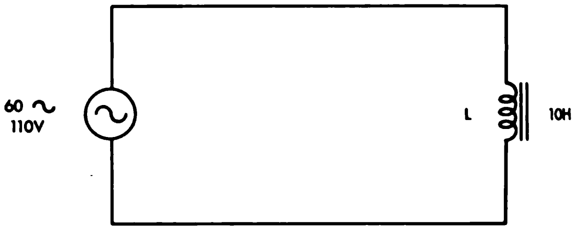
Suppose that you want to find the inductive reactance of a 10-henry coil when the frequency of the applied alternating current flowing through the coil is 60 cps.

Solution:

Substituting in the equation for inductive reactance,

$$X_L = 2\pi f L$$
$$= 6.28 \times 60 \times 10$$
$$= 3{,}768 \text{ ohms}$$

If this 10-henry coil is connected across a 110-volt, 60-cycle source of alternating current, you can calculate the current that will flow in the circuit by the formula,



Finding Inductive Reactance

$$I = \frac{E}{X_L}$$

where I is the current, E is the applied voltage, and X_L is the inductive reactance of the coil. Substituting the above values for E and X_L in this formula,

$$I = \frac{110}{3{,}768}$$
$$I = .029 \text{ ampere}$$

INDUCTORS CONNECTED IN SERIES

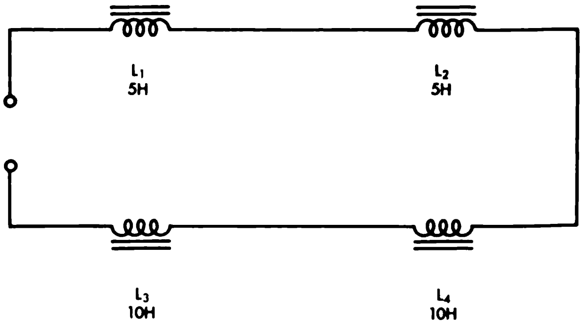
When two or more coils are connected in series, and there is no mutual inductance between them, the total inductance is equal to the sum of the individual inductances (assuming pure inductances). This is expressed by the formula,

$$L_t = L_1 + L_2 + L_3 + L_4$$

where L_t is the total inductance in henries and L_1, L_2, L_3, and L_4 are the individual inductances in henries.

Example 1

Suppose that you want to find the total inductance of a circuit when two 5-henry coils and two 10-henry coils are connected in series.



Illustrative Example

Solution:

Substituting the individual inductance values in the above formula,

$$L_t = 5+5+10+10$$
$$= 30 \text{ henries}$$

The total inductive reactance of a series circuit, consisting of two or more inductors, can be found by adding the individual inductive reactances. Expressed as a formula,

$$X_{L_t} = X_{L_1} + X_{L_2} + X_{L_3} + X_{L_4}$$

Example 2

Find the total inductive reactance of the inductors in the above example when the frequency of the applied voltage is 60 cps.

Solution:

Find the individual inductive reactances.

$$X_{L_1} = 2\pi f L$$
$$= 6.28 \times 60 \times 5$$
$$= 1{,}884 \text{ ohms}$$

Since L_2 and L_1 are equal, X_{L_2} also equals 1,884 ohms.

$$X_{L_3} = 2\pi f L$$
$$= 6.28 \times 60 \times 10$$
$$= 3{,}768$$

Since L_3 and L_4 are equal, X_{L_4} also equals 3,768 ohms.

Substituting in the formula for total inductive reactance, we see

$$X_{L_t} = 1{,}884 + 1{,}884 + 3{,}768 + 3{,}768$$
$$= 11{,}304 \text{ ohms}$$

The solution of this problem can be simplified since you know the value of total inductance.

$$X_{L_t} = 2\pi f L_t$$
$$= 6.28 \times 60 \times 30$$
$$= 11{,}304 \text{ ohms}$$

INDUCTORS CONNECTED IN PARALLEL

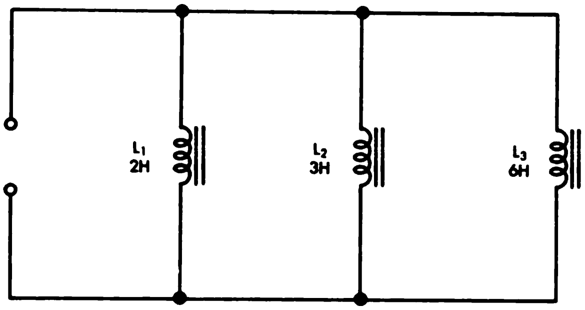
When a number of inductors are connected in parallel, and there is no mutual inductance between them, the total inductance is found by taking the reciprocal of the sum of the reciprocals (the same method that is used for finding the total resistance of two or more resistors connected in parallel). (Assume that we are again dealing with pure inductances.) This formula is

$$L_t = \frac{1}{\frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3}}$$

where L_t is the total inductance and L_1, L_2, and L_3 are the individual inductances.

Example 1

Three coils are connected together to make up a parallel inductive circuit. Their individual values are 2 henries, 3 henries, and 6 henries. Find the total inductance.



Illustrative Example 2

Solution:

Substituting these inductance values in the above formula for total inductance,

$$L_t = \frac{1}{\frac{1}{2} + \frac{1}{3} + \frac{1}{6}}$$

Finding the least common denominator of the terms in the denominator and solving,

$$L_t = \frac{1}{\frac{3}{6} + \frac{2}{6} + \frac{1}{6}}$$
$$= \frac{1}{\frac{6}{6}}$$
$$= 1 \text{ henry}$$

The total inductive reactance of a parallel circuit, consisting of two or more inductors, can be found by taking the reciprocal of the sum of the reciprocals of the individual inductive reactances. Expressed as a formula,

$$X_{L_t} = \frac{1}{\frac{1}{X_{L_1}} + \frac{1}{X_{L_2}} + \frac{1}{X_{L_3}}}$$

Example 2

Find the total inductive reactance of the parallel circuit in *Example 1* at an applied frequency of 60 cps.

Solution:

Find the individual inductive reactances:

$$\begin{aligned} X_{L_1} &= 2\,\pi\,f\,L \\ &= 6.28 \times 60 \times 2 \\ &= 753.6 \text{ ohms} \\ X_{L_2} &= 2\,\pi\,f\,L \\ &= 6.28 \times 60 \times 3 \\ &= 1130.4 \text{ ohms} \\ X_{L_3} &= 2\,\pi\,f\,L \\ &= 6.28 \times 60 \times 6 \\ &= 2260.8 \text{ ohms} \end{aligned}$$

Substituting these values in the above formula for total inductive reactance,

$$X_{L_t} = \frac{1}{\frac{1}{753.6} + \frac{1}{1130.4} + \frac{1}{2260.8}}$$

Finding the least common denominator for the terms in the denominator, and solving, you will find that

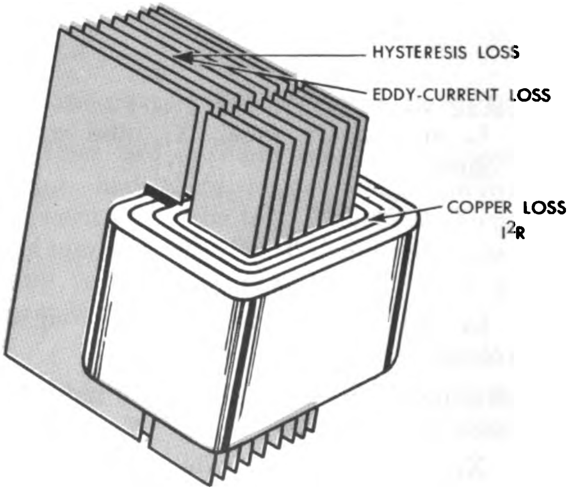
$$\begin{aligned} X_{L_t} &= \frac{1}{\frac{3}{2260.8} + \frac{2}{2260.8} + \frac{1}{2260.8}} \\ &= \frac{1}{\frac{6}{2260.8}} \\ &= \frac{2260.8}{6} \\ &= 376.8 \text{ ohms} \end{aligned}$$

The solution of this problem can be simplified since you know the value of the total inductance. Substituting the total inductance in the formula for inductive reactance,

$$\begin{aligned} X_L &= 6.28 \times 60 \times 1 \\ &= 376.8 \text{ ohms} \end{aligned}$$

POWER LOSSES IN INDUCTORS

Since inductors or coils consist of a number of turns of wire, and all wire has some resistance, all inductors have a certain amount of resistance. This resistance is small and is usually neglected in solving various types of AC circuit problems, because the reactance of the inductor is so much greater than the resistance that the resistance has a negligible effect on the current flow and the phase angle. However, since some inductors are designed to carry relatively large amounts of current, considerable power is dissipated in the inductor even though the amount of resistance of the winding is small. This is wasted power and is called a *copper loss*. The copper loss of an inductor can be calculated by multiplying the square of the current through the inductor by the resistance of the winding (I^2R). Typical inductors of this type are given maximum current ratings by the manufacturer, according to the length and diameter of the wire that makes up the winding. Keeping this figure in mind, you can usually prevent damage to an inductor, due to excessive current, by using an inductor whose current rating is sufficiently high for a particular circuit.



Power Losses in Inductors

Iron-core coils have *iron losses* in addition to copper losses. Iron losses are divided into hysteresis losses and eddy-current losses. *Histeresis losses* are due to energy that is consumed in reversing the magnetic field of the core each time the direction of current flow through the winding changes. In the study of magnetism, you learned that the axes of the molecules of a magnetized piece of iron are aligned so that they are all parallel with the magnetic field. In an iron-core coil, these axes are turned back and forth at a rate determined by the frequency of the current flowing through the winding. *Eddy-current losses* are due to currents that are induced in the iron core by the magnetic fields around the turns of the coil. These currents are called eddy currents. They flow back and forth in the iron and dissipate energy in the form of heat in the iron core. Since this energy cannot be returned to the electrical circuit, it is lost or wasted energy.

Iron cores are seldom solid pieces of iron since the eddy-current losses in a solid iron core would be excessive. A common type of iron core that is used in power-frequency and audio-frequency coils is the *laminated core.* The laminated core consists of thin sheets of soft iron, called laminations. These laminations are insulated from each other by coatings of shellac or varnish and are held together by the frame of the coil or by small bolts. The number of laminations required in the core of a certain coil is determined by the amount of current that the coil must be able to carry without becoming saturated. This method of core construction limits the lengths of the eddy-current paths to the thickness of the laminations (from .01 to .05 centimeter) and thereby limits the eddy-current losses.

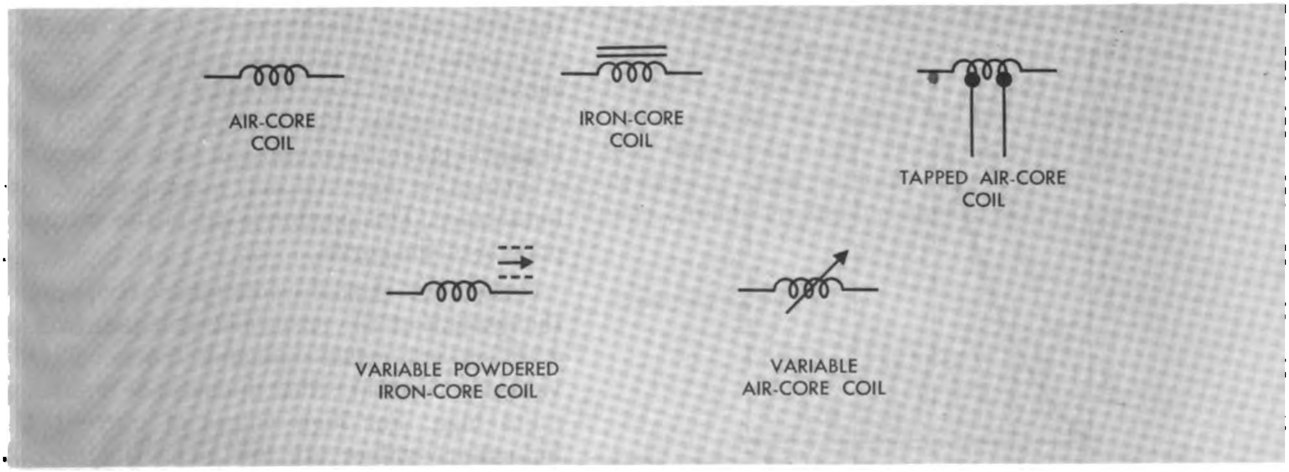
A type of core that is used in coils designed for frequencies above the audio frequency range is the *powdered-iron core.* This type of core is made by mixing powdered iron particles in an insulating fluid and hardening the mixture into the proper shape. This type of core has even fewer losses than the laminated core since the eddy-current paths are restricted to the individual particles of powdered iron.

TYPICAL INDUCTORS

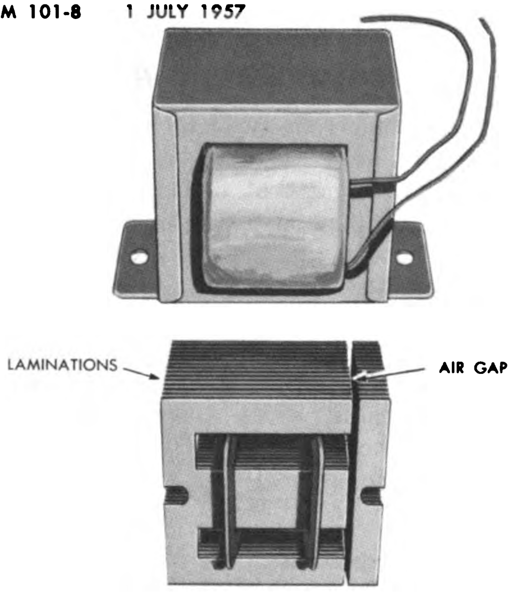
There are numerous types of inductors used in communications equipment, differing in physical size, amount of inductance, type of core, type and method of winding, number of turns, size of wire, and type of insulation. The symbols for these various types of coils are shown in the illustration below. The simplest way to classify typical inductors for a general discussion is into three groups according to the frequencies at which they are designed to operate. These frequencies are: power frequencies, audio frequencies, and radio frequencies.

For Power Frequencies

FILTER CHOKES. Nearly all pieces of communications equipment require a source of direct current as well as alternating current. When the available power source is alternating current, some of the alternating current must be changed to direct current. In the process of converting from alternating current to direct current, it is necessary to rectify the alternating current, resulting in pulses of current that flow in the same direction. These pulses must be smoothed out to obtain a constant value of direct current. Since an inductor stores energy when the current flowing through it is increasing and returns the energy to the circuit when the current is decreasing, it is used in smoothing out these pulses. Inductors used for this purpose are called *power chokes* or *filter chokes* since they smooth

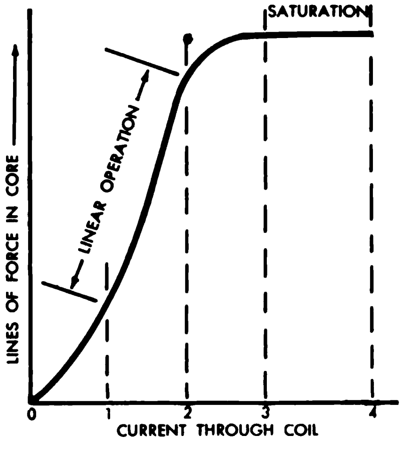


Electrical Symbols for Inductors



Laminated Iron-core Coil

out or choke out the variations in current. A choke coil has a laminated iron core, since a large amount of inductance is required while the losses must be kept to a minimum. The size of wire used in the winding must be large enough so that the coil will carry the required amount of current without introducing too great a voltage drop and without becoming so hot that the insulation becomes damaged. The insulation on the windings must be sufficient so that an arc does not jump across the windings due to the relatively large amount of voltage existing between the turns. The inductance of typical filter chokes ranges from about 2 henries to 30 henries, while their current carrying capacities range from about 50 milliamperes (50 thousandths of an ampere) to about 5 amperes.



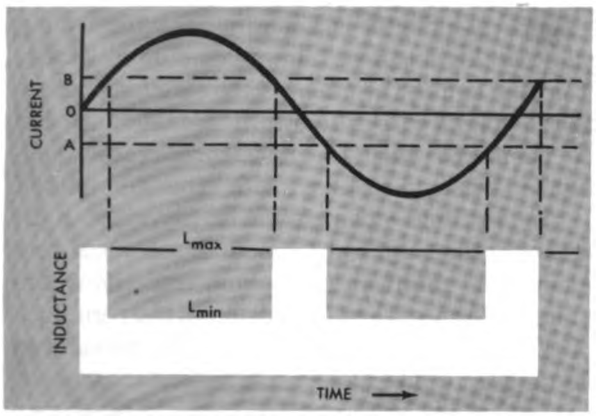
Magnetization Curve for an Iron-core Coil

Nearly all filter chokes have shell-type cores similar to the one shown in the illustration. This core is not a complete magnetic path, since the path is broken by an air gap. The air gap increases the reluctance of the total magnetic path and thereby minimizes the tendency for the choke to become saturated.

SATURABLE REACTORS. Another type of inductor depends upon the amount of current flowing through it for its operation. In order to explain the effect of current upon the inductance of a coil, it is first necessary to discuss magnetic saturation. According to the molecular theory of magnetism, a magnet consists of numerous molecular magnets. In other words, the electron orbits around the center of the atoms are all aligned in the same direction. Since each revolving electron can be considered as a current flow, each atom has a magnetic field. If all of these fields are in the same direction, the whole piece of material is magnetic. In a nonmagnetic material, the molecular magnets are arranged in a random fashion so that the total magnetic field is zero. When all of these molecular magnets are aligned in the same direction in an iron-core coil carrying current, additional current through the winding cannot increase the strength of the magnetic field. The core is then said to be *saturated.*

On the Magnetization Curve for an Iron-Core Coil, the number of lines of force start increasing as current starts to increase from zero. Between points 1 and 2, there is an equal increase in the number of lines of force for equal amounts of increase in current. This portion of the curve is therefore called the linear portion. From point 2 to point 3, the increase in lines of force becomes less and less for equal amounts of increase in current. Between points 3 and 4 there is no change in the number of lines of force with a change in current and, consequently, little inductance. This portion is called the saturation portion of the curve.

Some filter chokes are designed to operate on the portion of the curve between points 1 and 4. They are called *swinging chokes.* For low values of current, their inductance values are quite high. For high values of current their inductance values are low since the core has become saturated. The air gap in the swinging choke is much smaller than in the ordinary



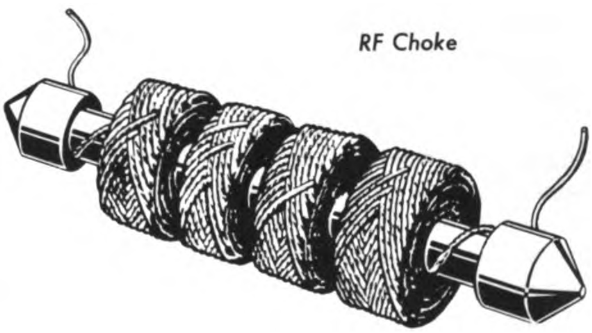
Current and Inductance in Saturable Reactor

filter choke and, consequently, the total reluctances or opposition to the magnetic lines of force through the core is small. This permits the choke to become saturated for high values of current. Swinging chokes have special applications in circuits where the direct-current variations are quite large since it has been found that they give better direct voltage regulation with varying amounts of direct current — that is, the voltage remains nearly constant even though the amount of current varies over a wide range.

Another type of saturable reactor is designed so that an abrupt change in inductance occurs when the current exceeds the saturation point or falls below the saturation point. This type has special applications in circuits that are designed to change the shape of the waveform from a sine wave to a square wave as shown in the illustration above. When the amount of current is between the limits of B and A, the reactor is not saturated and it has a maximum amount of inductance. Since it has a maximum amount of inductance, the inductive reactance is also maximum and a large voltage drop is developed across the reactor. When the current rises above amount B or falls below amount A, the reactor becomes saturated and the inductance falls to a minimum value. Therefore, the inductive reactance also falls to a minimum value and, consequently, only a small amount of voltage is developed across the reactor. The square wave is discussed further in chapter 30.

For Audio Frequencies

Inductors designed for audio frequencies are called audio chokes. They have two general purposes. One purpose of the audio choke is similar to the purpose of the power choke; that is, it is used to smooth out variations in current occurring at audio frequencies (16 to 16,000 cps). The other purpose of the audio choke is to produce a large audio voltage drop or a large change in audio voltage for a relatively small change in audio current. The inductance of the audio choke must be high in both cases. Typical values of audio chokes range from several henries to 500 henries. These chokes are similar in appearance to power chokes, except that they are usually somewhat smaller in physical size. The size of wire used for the windings is relatively small since they do not have to carry as much current as power chokes. Current ratings of these chokes range from about 5 to 100 milliamperes. Nearly all audio chokes have laminated iron cores that are similar to the cores in power chokes.

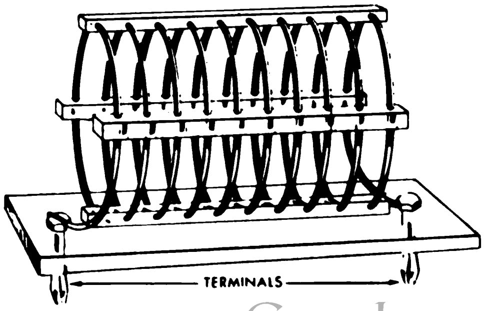


RF Choke

For Radio Frequencies

There are two types of inductors designed for use at radio frequencies. One type is the radio-frequency choke shown in the illustration above. This type serves two purposes, similar to the two purposes of audio chokes. It is used to smooth out variations in current at radio frequencies. The radio-frequency choke is also used to develop a large radio-frequency voltage drop from a relatively small change in radio-frequency

Fixed Tuning Coil



current. Since inductive reactance varies directly with frequency, and radio frequencies are much higher than audio frequencies, the amount of inductance of a radio-frequency choke is quite small in comparison to the inductance of an audio choke. A typical inductance rating of a radio-frequency choke is 2.5 millihenries. Since there is very little energy in most radio-frequency circuits, the losses must be held to a minimum. For this reason, this type of choke usually has an air core, since air-core losses are negligible.

The other type of inductor designed for use at radio frequencies is used in tuning circuits which can be made to select a desired radio frequency out of a number of frequencies or to determine the frequency of a radio-frequency generator. These coils have very small amounts of inductance (2 to 300 microhenries or millionths of a henry). In respect to the other two general types of inductors, these coils have very few turns and either have air cores or powdered-iron cores.

Some radio-frequency tuning coils have fixed values of inductance and some are designed so that the amount of inductance is variable. Variable inductors are used in circuits that can be made to select various frequencies, one at a time, out of a number of frequencies. These are continuously variable from a minimum to a maximum amount of inductance. One common type of variable inductor has a powdered-iron core that can be moved in and out of the center of the coil by means of a threaded brass screw that is imbedded in the powdered-iron core. Other inductors used in tuning circuits have fixed taps at various points, so that various fixed amounts of inductance can be selected by some mechanical arrangement.

TESTING TYPICAL INDUCTORS

An ohmmeter is very convenient in making simple checks in respect to different types of inductors. It is an instrument designed to measure resistance. A simple ohmmeter circuit consists of a battery, a variable resistor (zero adjust), and a sensitive current meter, connected in series. When the test probes of the ohmmeter are placed on the two terminals of a coil, the battery current of the instrument must also flow through the coil, and the added resistance of the coil is indicated by the meter.

Since an inductor is a continuous wire, wound on a core, it has a certain amount of resistance from one end to the other. If the exact resistance of the coil of wire is known, this resistance can be compared with the resistance reading obtained when measured with the ohmmeter. If the readings compare favorably, it can be assumed that the coil is good. If no resistance reading is obtained (the ohmmeter shows infinite resistance when placed across the terminals of the coil), the coil is open. (The term *open* is applied to any circuit component that measures infinite resistance when it should measure a definite amount of resistance.) An open coil is usually the result of excessive heating of the coil due to excessive current flow, or to vibration.

If the reading is more than about 20% below the rated value, a number of turns of the coil are probably shorted together, reducing the amount of resistance (and also inductance) of the coil. (The term *short* is applied to any circuit component, or part of a component, that measures a very low or zero value of resistance when its normal resistance reading should be a much higher value.) Shorted turns may be caused by excessive heating due to too much current flowing through the coil. They may also be caused by *chaffing*. As the current varies through the coil, the individual turns are attracted to each other with a varying force. This may cause them to rub against each other, causing the insulation to wear off or chaff.

If the measured resistance across the terminals of the coil is very low or zero, the whole coil is shorted. This may be due to the two terminals being shorted directly to the core in an iron-core coil. The resistance of the iron-core is very low.

In the case of an iron-core coil, it is always well to check the resistance from one or both of the terminals (one at a time) to the core itself, to determine whether or not the winding has been shorted to the core. This also is a common occurrence when a coil is overheated due to excessive current flow. When this occurs, the insulation becomes burned off at a particular point and allows the wire to touch the core.

It is usually impractical to repair iron-core coils, such as filter chokes, when any of the above mentioned failures have occurred, and they must therefore be replaced. However, some radio-frequency coils, especially the space-wound type, can be repaired.

Chapter 14

The Basic Capacitor

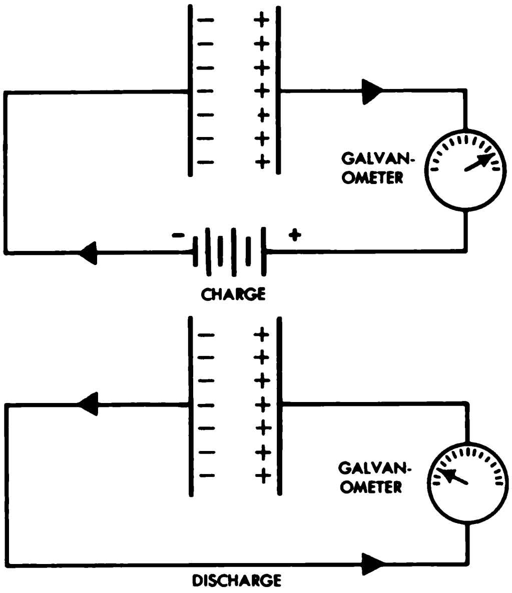
The second type of reactor is the capacitor. This chapter deals with capacitors, their physical and electrical characteristics, and their effects in AC circuits. Practical methods of checking capacitors are also described.

THE BASIC CAPACITOR

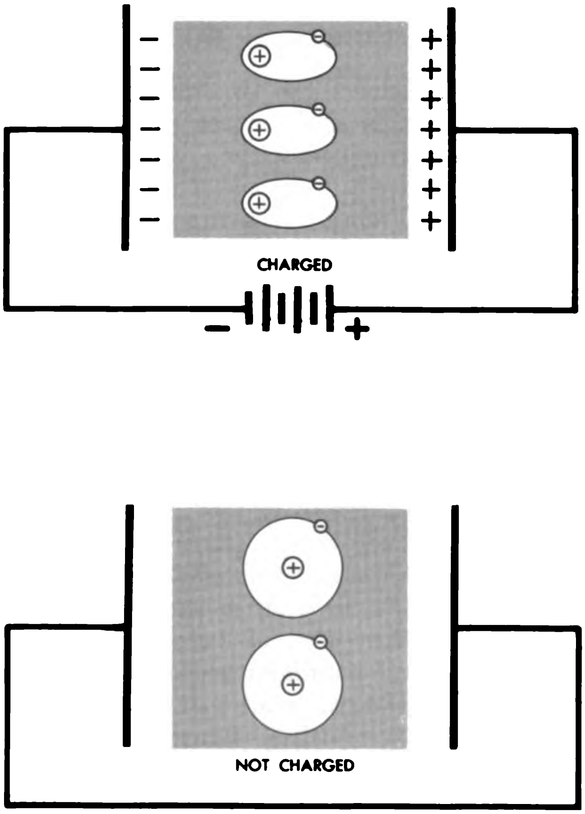
If you place two sheets of aluminum or copper near each other, but not so close that they are touching, and connect the sheets to the opposite terminals of a battery as shown on page 100, a momentary current will flow in the circuit. This momentary current can be observed if you insert a galvanometer in series with one of the battery terminals. According to the basic principles that you learned in the study of direct current, this would be called an open circuit, and yet a momentary current flows. Actually, this same phenomenon occurs when two wires are connected to the terminals of a battery and the ends of the wires are held near each other. However, since current flows for only an instant and then ceases, both of the above examples are effectively open circuits for direct current.

If you disconnect the two plates from the battery terminals and connect them together through a galvanometer, you can observe a momentary current flow again, but in the opposite direction. From this experiment it is apparent that the two plates store energy when connected to a voltage source and release the energy when they are connected together. It can be said that they have a capacity for storing electric energy and together they are called a *capacitor*. The words capacitor and condenser are used interchangeably. Inductors store energy in their magnetic fields, while capacitors store energy in their electric fields.

The material between the plates of a capacitor is called the *dielectric*. Some of the various dielectric materials used in capacitors are air, mica, and oil. Whatever the dielectric material, it is made up of molecules and atoms. Atoms consist, basically, of electrons that revolve around a positive nucleus. When you connect the capacitor to a source of voltage, the electron orbits in the dielectric material are distorted away from the negative terminal and toward the positive terminal, since like charges repel



Basic Capacitor



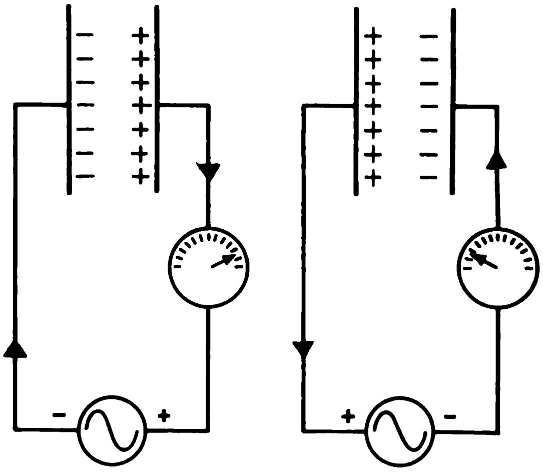
Charge and Discharge of a Capacitor Dielectric

and unlike charges attract. The plate connected to the negative terminal receives an excess of electrons and is therefore given a negative charge, and the plate connected to the positive terminal loses some of its electrons and therefore becomes positively charged (diagram A). When you disconnect the capacitor from the voltage source, the plates hold their charges and the electron orbits remain distorted. When you connect the plates together, electrons flow from the plate with the excess electrons until the charges on the two plates are equalized and the electron orbits in the dielectric material return to their normal shape. The capacitor is then said to be discharged (diagram B).

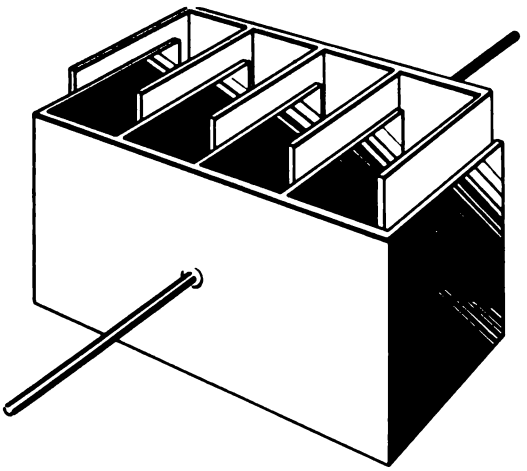
When an alternating voltage is applied across a capacitor as shown in the illustration below, the capacitor charges in one direction, discharges, and charges in the other direction, as the polarity of the alternating voltage changes. Even though electrons do not actually pass through the dielectric from one plate to the other and back, there is a continuous flow of alternating current when a capacitor is connected in an alternating-current circuit. It can be said that a capacitor blocks the flow of direct current and is a conductor for alternating current.

FACTORS AFFECTING CAPACITANCE

The capacity for storing energy in an electric field is commonly called *capacitance*. The amount



AC Applied to a Capacitor



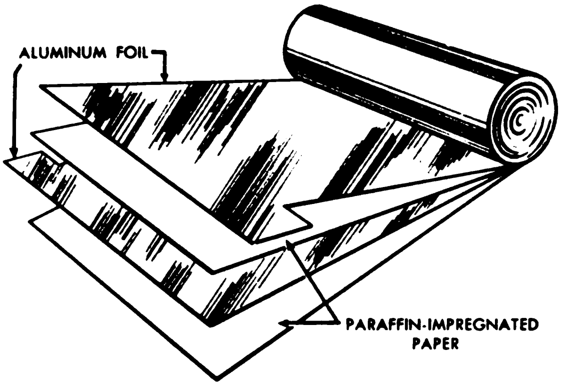
Multi-plate Capacitor

of capacitance of a capacitor depends upon the *area of the plates*, *the distance between the plates*, and the *type of dielectric material* that is used in a particular capacitor.

Area of the Plates

It was mentioned that charging current would flow when the ends of two wires that were connected to the terminals of a battery were held near each other. The amount of this current would be so small that it could not be detected by means of a galvanometer connected in the circuit. On the other hand, the charging current of the two plates would give a very noticeable deflection. The reason for this effect lies in the difference between the area of the ends of the wires and the area of the plates. *The larger the area of the plates of a capacitor, the greater the capacitance.* Some typical capacitors consist of two plates, similar to the plates of the basic capacitor in the previous discussion. Other capacitors consist of a number of plates; every other one being connected together to form effectively a single plate of large area (multi-plate capacitor). When a number of plates are connected together in this manner, the total area is equal to the sum of the areas of adjacent plates. Considering both sides of each plate, the total area of the capacitor illustrated above is equal to the area of one plate times 9.

Another type of capacitor is constructed of two long, narrow sheets of aluminum foil that are separated from each other by a strip of paraffin-impregnated paper which acts as the dielectric. These long strips are wound into a cylindrical tube and make up what is called a tubular paper capacitor. The area of the long strips of aluminum foil is the area of the plates.



Tubular Paper Capacitor

Distance between the Plates

The distance between the plates and the thickness of the dielectric are one and the same. *The greater the distance between the plates, the less the capacitance.* This is true because the greater the distance between two charged particles, the less the force that exists between them.

Type of Dielectric Material

The electric orbits in some materials can be distorted by less force than the electron orbits in other materials. This gives rise to a value assigned to dielectric materials called the *dielectric constant.* It requires much less force to distort the electron orbits in a piece of mica, for example, than it does to distort the electron orbits in air, by the same amount. Air is taken as a reference and given a dielectric constant of 1. In reference to air, it requires only about one-sixth as much force to distort the electron orbits in mica; therefore the dielectric constant of mica is 6. In other words, if a mica-dielectric capacitor and an air-dielectric capacitor of the same size are connected to equal voltage sources, the mica-dielectric capacitor will store six times more energy than the other. *The greater the dielectric constant, the greater the capacitance.*

FORMULAS FOR CAPACITANCE

The factors that determine the capacitance of a capacitor can be summed up by the formula, $C = 0.2248\frac{AK}{t}$

where C is the capacitance in *micromicro-farads,* K is the dielectric constant, A is the area in square inches, and t is the distance between the plates in inches. The areas of the opposite plates of a capacitor are equal; therefore, the area of one plate is A in the above formula.

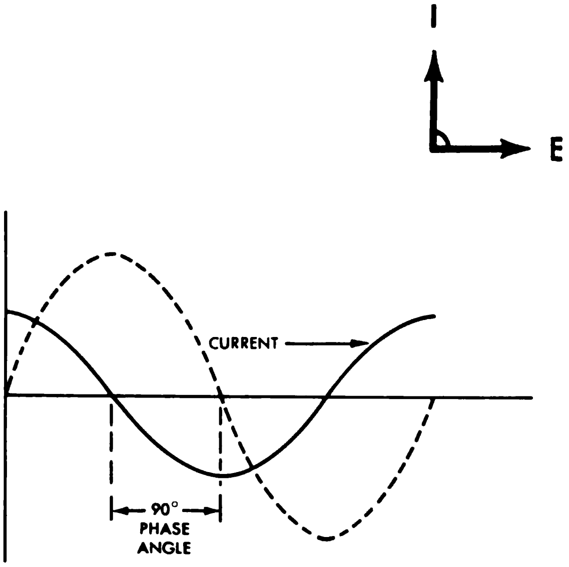
The unit of capacitance (the farad) is the amount of capacitance which will store one coulomb at a potential of one volt. A coulomb is a quantity of charge equal to 6.28×10^{18} electrons. The number of coulombs stored by a capacitor is equal to the applied voltage times the capacitance in farads. This is expressed by the formula, $Q = EC$.

In practical circuits, the values of capacitors are very small in respect to the farad and are therefore expressed in microfarads (mf) and micromicrofarads (mmf).

PHASE RELATIONSHIPS IN A CAPACITOR

A capacitor has a certain amount of resistance since the material of which the plates are made has a small amount of resistance and the dielectric material has a certain high value of resistance. The resistance of the plates is effectively in series with the capacitor and the resistance of the dielectric is effectively in parallel, or across, the plates of the capacitor. Since the total resistance across the plates of the capacitor is high in comparison to the capacitive reactance, the resistance can be neglected in most discussions. The capacitor is then said to be a *pure capacitance.*

When the voltage applied across a capacitor is increasing, current will flow towards the plate connected to the side of the power source that is negative and away from the plate connected to the side that is positive. As the voltage reaches its maximum value and stops increasing, there will be no current flowing in the circuit. In other words, the applied voltage is at its maximum and the current in the circuit is at zero so that a phase angle of 90° exists between the voltage and current. When the applied voltage starts decreasing from its maximum positive value, the voltage across the capacitor must decrease and therefore the capacitor starts to discharge. The reversal in current direction in the capacitive circuit occurs 90° before the reversal in polarity across the capacitor; therefore, *the current leads the voltage by 90° in a purely capacitive circuit.*



Current Leads Voltage by 90° in a Capacitor

CAPACITIVE REACTANCE

The opposition of a capacitor to the flow of alternating current is called *capacitive reactance.* The amount of capacitive reactance depends upon the capacitance of the capacitor, the frequency of the applied alternating voltage, and the constant $\frac{1}{2\pi}$ or .159. Expressed as a formula,

$$X_C = \frac{1}{2\pi f C}$$

where X_C is the capacitive reactance in ohms, f is the frequency of the applied voltage in cycles per second, and C is the capacitance of the capacitor in farads. The capacitive reactance is said to vary inversely as the frequency and capacitance — that is, the higher the frequency the less the capacitive reactance; and the greater the amount of capacity, the less the capacitive reactance.

Example

Find the capacitive reactance of a .01 microfarad capacitor at a frequency of 60 cps.

Solution:

$$X_C = \frac{1}{2\pi f C}$$

$$= \frac{0.159}{60 \times .01 \times 10^{-6}}$$

(0.1 microfarad is expressed as $.01 \times 10^{-6}$)

$$X_C = \frac{0.159}{0.6 \times 10^{-6}}$$

$$= 0.265 \times 10^{6}$$

$$= 265{,}000 \text{ ohms}$$

If this .01-microfarad capacitor is connected across a 110-volt, 60-cycle source of alternating current, you can calculate the current that will flow in the circuit by the formula,

$$I = \frac{E}{X_C}$$

Substituting the above values for E and X_C,

$$I = \frac{110}{265{,}000}$$

$$= .0004 \text{ ampere}$$

$$= 0.4 \text{ milliampere}$$

CAPACITORS CONNECTED IN SERIES

When two or more capacitors are connected in series, the total capacitance is equal to the reciprocal of the sum of the reciprocals of the individual capacitances. This is similar to the formula for finding total inductance when inductors are connected in parallel.

Expressed as a formula,

$$C_t = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}}$$

where C_t is the total capacitance, and C_1, C_2, and C_3 are the individual capacitances.

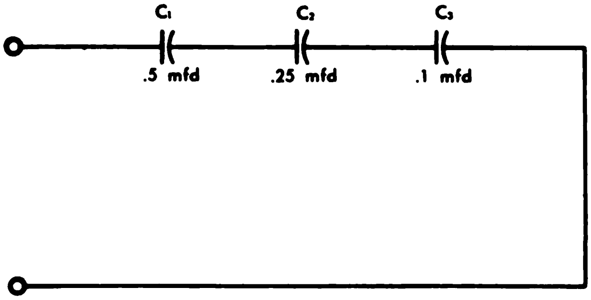
Note that if three capacitors of equal size are connected in series, the effect on the total capacitance is the same as if the distance between the plates of one of the capacitors were increased to three times its original distance. As a result the total capacitance becomes one-third the capacitance of one capacitor.

Example 1

Find the total capacitance of a circuit consisting of a 0.5-mfd, a 0.25-mfd, and a 0.1-mfd capacitor connected in series.

Solution:

Substituting the above values in the formula,



Illustrative Examples

$$C_t = \frac{1}{\frac{1}{0.5 \times 10^{-6}} + \frac{1}{0.25 \times 10^{-6}} + \frac{1}{0.1 \times 10^{-6}}}$$

Finding the lowest common denominator of the terms in the denominator and solving,

$$C_t = \frac{1}{\frac{1 + 2 + 5}{0.5 \times 10^{-6}}}$$

$$= \frac{1}{\frac{8}{0.5 \times 10^{-6}}}$$

$$= .0625 \times 10^{-6}$$

$$= .0625 \text{ mfd}$$

The total capacitive reactance of capacitors connected in series is equal to the sum of the individual capacitive reactances,

$$X_{C_t} = X_{C_1} + X_{C_2} + X_{C_3}$$

Example 2

Find the total capacitive reactance of the three capacitors in the above example when the frequency is 60 cps.

Solution:

$$X_{C_1} = \frac{0.159}{60 \times 0.5 \times 10^{-6}} = 5{,}300 \text{ ohms}$$

$$X_{C_2} = \frac{0.159}{60 \times 0.25 \times 10^{-6}} = 10{,}600 \text{ ohms}$$

$$X_{C_3} = \frac{0.159}{60 \times 0.1 \times 10^{-6}} = 26{,}500 \text{ ohms}$$

$$X_{C_t} = 5{,}300 + 10{,}600 + 26{,}500$$

$$= 42{,}400 \text{ ohms}$$

The solution of this problem can be simplified since you know the value of the total capacitance.

$$X_{C_t} = \frac{0.159}{60 \times .0625 \times 10^{-6}}$$

$$= \frac{0.159}{3.75 \times 10^{-6}}$$

$$= .0424 \times 10^{6} = 42{,}400 \text{ ohms}$$

CAPACITORS CONNECTED IN PARALLEL

When two or more capacitors are connected in parallel, the total capacitance is equal to the sum of the individual capacitances. This is expressed by the formula,

$$C_t = C_1 + C_2 + C_3$$

where C_t is the total capacitance, and C_1, C_2, and C_3 are the individual capacitances.

Example

Find the total capacitance of the parallel circuit shown below.

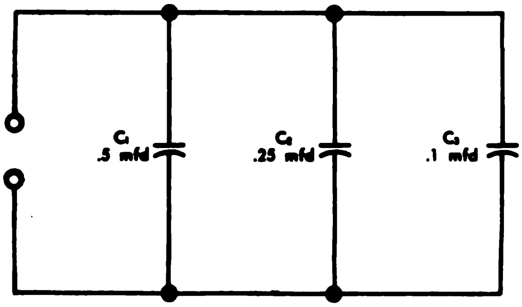


Illustration Example

Solution:

$$\begin{aligned} C_t &= 0.5 + 0.25 + 0.1 \\ &= 0.85 \text{ mfd} \end{aligned}$$

The total capacitive reactance of two or more capacitors connected in parallel is equal to the reciprocal of the sum of the reciprocals of the individual capacitive reactances. This is expressed by the formula,

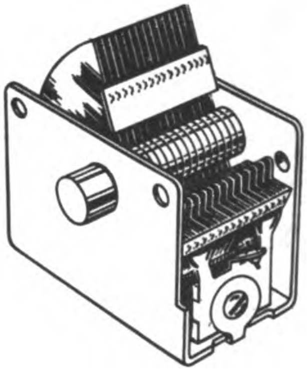
$$X_{C_t} = \frac{1}{\frac{1}{X_{C_1}} + \frac{1}{X_{C_2}} + \frac{1}{X_{C_3}}}$$

The solution of this particular problem can be simplified since you already know the total capacitance of the circuit. At 60 cps,

$$\begin{aligned} X_{C_t} &= \frac{0.159}{60 \times 0.85 \times 10^{-6}} \\ &= \frac{0.159}{51 \times 10^{-6}} \\ &= .0031176 \times 10^{6} \text{ or } 3{,}118 \text{ ohms} \end{aligned}$$

POWER LOSSES IN CAPACITORS

If it were possible to construct a perfect capacitor, it would store a certain amount of energy when charged, and deliver all of that energy back to the circuit when discharged. However, it is not possible to construct a perfect capacitor and some power is lost or dissipated in any capacitor. Capacitor losses are classified as *resistance losses*, *leakage losses*, *dielectric-absorption losses*, and *dielectric-hysteresis losses.*



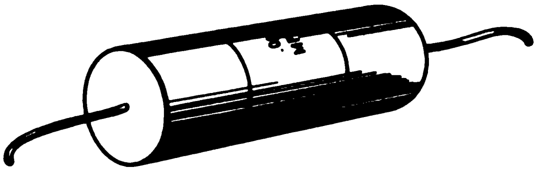
Variable Capacitor

Resistance Losses

Resistance losses are due to the resistance of the wires that are connected to the plates and to the resistance of the plates themselves. The resistance losses equal the square of the current times the resistance (I^2R). These losses represent only a small amount of the total loss of energy in a capacitor.

Leakage Losses

Leakage losses are due to small amounts of current that flow through the dielectric material from one plate to the other. Air dielectric and mica dielectric capacitors have very small leakage losses, while other types have considerable leakage losses. These losses result in heat being



Tubular Paper Capacitor

developed in the dielectric material. Excessive leakage losses may cause enough heat to be generated in the capacitor to destroy the insulating qualities of the dielectric and thereby minimize or even eliminate the capacitance of the capacitor.

Dielectric-Absorption Losses

Some dielectric materials, such as paraffin-impregnated paper, seem to absorb charges as the capacitor is being charged. These absorbed charges are not given up as the capacitor is discharged and consequently constitute an energy loss. These losses are minimized in high-quality capacitors by making the paraffined paper of uniform thickness and free from foreign substances and particles.

Dielectric-Hysteresis Losses

Dielectric hysteresis is similar to the hysteresis in the core of an inductor. It requires energy to reverse the polarity of the electric field existing across the plates of the capacitor, the same as it requires energy to reverse the magnetic field in the core of an inductor. This form of loss results in heat, the amount of which is determined by the frequency of the alternating voltage across the capacitor. These losses become quite noticeable at the higher radio frequencies.

Voltage Ratings

Another important characteristic of a capacitor is its voltage rating. This may be expressed as either a DC working voltage, peak voltage, or surge voltage rating. Working voltage is the maximum DC voltage that the capacitor can withstand under continuous operation at normal operating temperatures. Peak voltage is the maximum peak value of AC voltage that the capacitor can withstand under continuous operation. Surge voltage is the maximum voltage a capacitor can withstand under certain conditions for five minutes. This last rating is valuable information when you are analyzing circuits that are subject to periodic voltage surges.

TYPICAL CAPACITORS

The numerous types of capacitors that are used in communications and radar equipment can be conveniently classified according to the types of dielectric materials used. There are five dielectric materials in common use. They are *air*, *paraffin-impregnated paper*, *mica*, *oil*, and *oxides*.

Capacitors with Air Dielectric

Capacitors that are used for tuning purposes (selecting stations or frequencies) use air as the dielectric. These capacitors are called variable capacitors see page 104. One set of plates is secured to the frame of the capacitor by insulators, and the other set of plates is fastened to a shaft that can be turned since it is suspended between bearings. The variable plates mesh with the fixed plates so that one variable plate passes between two fixed plates. A typical capacitor of this type that is used in a broadcast receiver has a minimum capacitance of 35 mmf and a maximum capacitance of 365 mmf.

Capacitors with Paper Dielectric

Capacitors that have a dielectric material made up of paper that has been dipped in molten paraffin are called paraffin-impregnated paper dielectric capacitors. These capacitors have a relatively high dielectric loss, especially at the higher radio frequencies. They are therefore more commonly used in audio-frequency and low-radio-frequency circuits. They are also called tubular paper capacitors as mentioned in the discussion of factors affecting capacitance.

Capacitors With Mica Dielectric

Capacitors having mica dielectrics are used in circuits that carry high radio-frequency currents. Mica has less dielectric loss than any of the other common types and, therefore, can be used without causing too much loss in circuits where the amount of available energy is very small. These capacitors are constructed of small rectangular plates separated by thin sheets of mica.

Capacitors with Oil Dielectric

Oil can withstand higher voltages than the other types of dielectric materials and, therefore, oil-dielectric capacitors are used in circuits that are subjected to especially high voltages such as those found in radio transmitters. Since oil can withstand higher voltages, a break-down of this type of capacitor (an arc jumps across the plates), will not damage the capacitor if the duration of the arc is not too long.

Electrolytic Capacitors

An electrolytic capacitor is constructed of aluminum plates and a chemical in liquid or paste form (ammonium borate or sodium phosphate). A direct voltage is applied across the terminals of this capacitor forming a thin film of insulating oxide on the positive plate. This oxide film acts as the dielectric. Since the film of oxide is very thin, this type of capacitor has a large amount of capacitance for a relatively small physical size. Typical values range from 8 microfarads to 200 microfarads. These capacitors cannot be used as conductors for AC since one plate must always remain positive in respect to the other; otherwise, the dielectric will be destroyed. However, they are commonly used in conjunction with a power (or filter) choke in smoothing out current variations in pulsating DC circuits.

TESTING TYPICAL CAPACITORS

It was mentioned at the end of the discussion on inductors that the ohmmeter may be used to check inductors. The ohmmeter can also be used to make several simple checks on capacitors. Small capacitors, such as those having mica dielectrics, can be checked for a short by measuring the resistance between the two capacitor terminals. If an infinite reading is obtained, the capacitor is probably good. If a reading from less than infinite to zero is obtained, the dielectric material has been damaged and the capacitor is no longer good. All capacitors have a voltage rating that is predetermined by the thickness and type of dielectric material. If this voltage rating is exceeded, the capacitor will probably arc over between the plates and burn or char the dielectric material. When this occurs, a small amount of carbon is formed which becomes a very good leakage path between the plates. This condition destroys the usefulness of the capacitor and it must be replaced.

Tubular paper capacitors that have a capacitance of .01 mfd or more will show a small momentary deflection on the ohmmeter when the ohmmeter is connected across the capacitor terminals. This shows that the capacitor receives a momentary charge from the ohmmeter battery. If the ohmmeter reading remains less than infinite, the paper dielectric has probably been burned or charred and the capacitor is of no further value.

Electrolytic capacitors will show a large deflection, when the terminals are connected to the ohmmeter test probes. In fact, if the polarity of the ohmmeter battery agrees with the polarity of the capacitor, the deflection will be almost the full width of the ohmmeter scale. This is due to the large amount of charging current that flows into the electrolytic capacitor. The reading should gradually rise as the capacitor becomes charged to at least 1 megohm (or slightly less if the polarity of the ohmmeter is reversed). If the reading remains at a value less than one megohm, the oxide film has probably been damaged and the capacitor is useless.

Chapter 15

Transformers

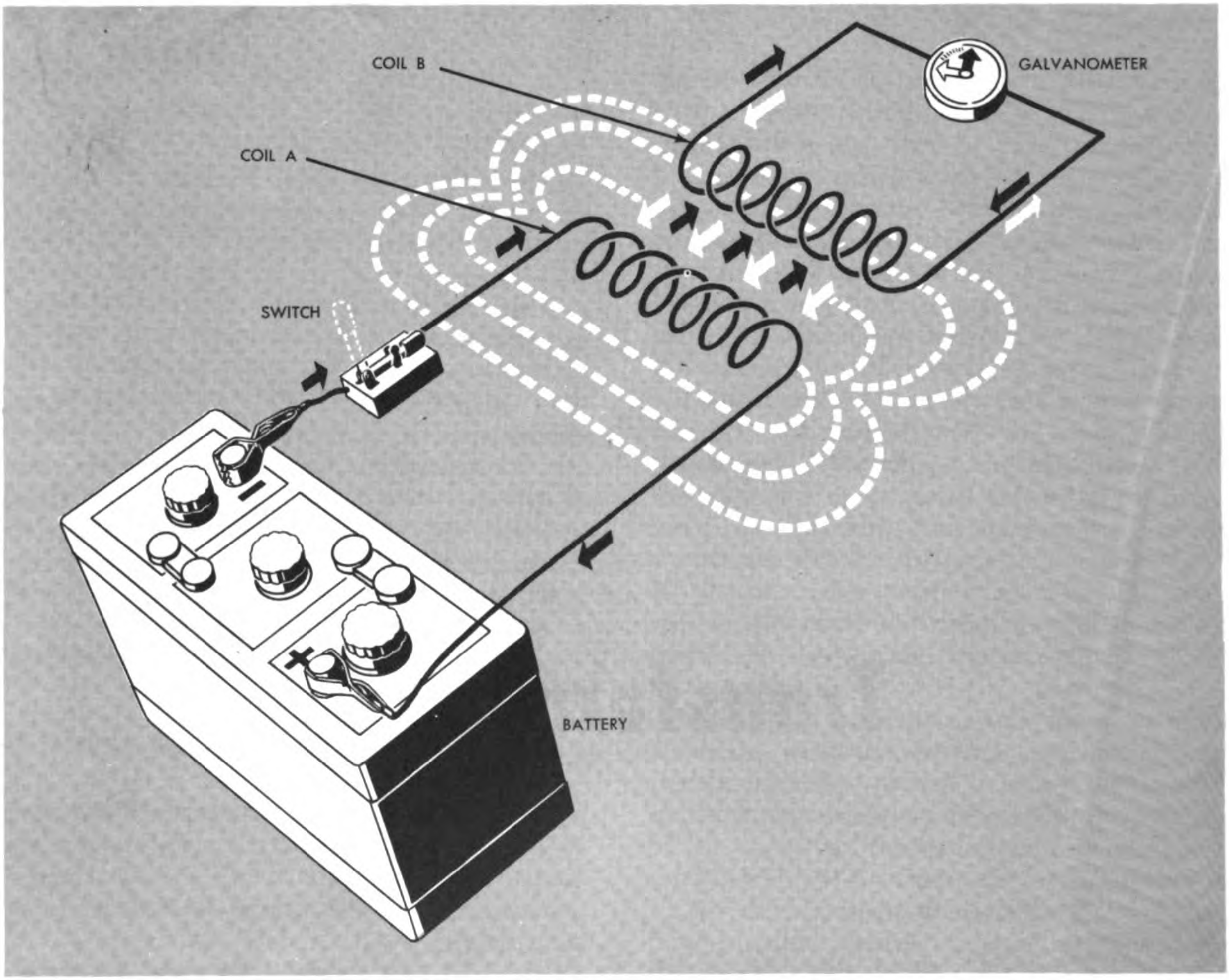
In the introduction to chapter 12, it was stated that an alternating voltage could be stepped up for the transmission of power over a long power line and stepped down again to the proper value for the consumer. It was explained that this procedure is used to minimize the loss of power along the transmission line.

Since the current is stepped down when the voltage is stepped up, less current flows over the power line than would flow if the power was transmitted at a voltage rating of 110 volts. This makes it possible to use smaller conductors for the transmission line which permits a great saving. In addition, the smaller lines have less weight and therefore the supporting towers or poles do not have to be so sturdy and strong. This also reduces the expense of the transmission line.

The device that steps up or steps down the AC voltage is called a *transformer*. Transformers have many applications in communications and other electronic circuits in addition to their applications in the transmission of electric power. Transformers are designed to step up or down voltages at various frequencies. This chapter includes discussions of various transformers as well as the theory of transformers.

TRANSFORMER ACTION

Note the two coils in the illustration on page 108. A galvanometer is connected in series with Coil B, and a battery and switch are connected in series with coil A. When you close the switch, current surges through coil A causing an expanding magnetic field. This magnetic field



Transformer Action

cuts the turns of coil B, causing a momentary deflection of the galvanometer pointer. This shows that electrical energy is momentarily passed or transferred from one coil to the other. This phenomenon sometimes is called *transformer action.*

As long as the switch remains closed, a constant amount of current flows through coil A and, therefore, the magnetic field around it does not move. Consequently, there is no current flow in coil B. It requires a moving magnetic field to induce a voltage. However, when you open the switch, current flow in Coil A suddenly ceases, and the magnetic field collapses. In collapsing, it again induces a voltage and causes a momentary current flow through the galvanometer. However, this time the field moves to the left, and the resulting current through coil B is in the opposite direction. A special type of switch that is used in a practical circuit of this type is called a *vibrator.* Vibrators are used to interrupt the DC path to the primary in automobile-radio power transformers. They are also used in some portable and special-purpose Air Force communications equipment.

When an alternating current is permitted to flow through coil A, the magnetic field alternately expands and collapses. This varying magnetic field causes a continuous alternating current to flow in coil B. Coil B is commonly called the *secondary* or the *secondary winding,* while coil A is called the *primary* or the *primary winding.*

MUTUAL INDUCTANCE

When the primary current of a transformer causes current to flow in the secondary, a magnetic field is also established around the secondary because of the secondary current. This field expands and cuts the turns of the primary. Consequently, a voltage is induced back into the primary. Since the primary induces a voltage in the secondary and the resulting secondary current causes a magnetic field that induces a voltage back into the primary, it is said that mutual induction occurs and that the two coils have *mutual inductance.* The unit of mutual inductance is the *henry,* the same as the unit of self inductance. *Two coils are said to have a mutual inductance of one henry when a change in current of one ampere per second in one coil produces a voltage of one volt across the second coil.*

In an ideal transformer, the amount of mutual inductance would be equal to the square root of the product of the primary inductance and the secondary inductance. Expressed as a formula,

$$M = \sqrt{L_p \times L_s}$$

The individual inductances of the primary and the secondary, L_p and L_s, depend upon the factors that determine the inductance of a single coil.

It is impossible to construct an ideal transformer, because some of the lines of force developed around the primary will not cut the secondary turns. (These lines of force make up the *leakage flux.*) This leads to a factor called the coefficient of coupling. If the number of lines of force cutting the secondary is divided by the total number of lines developed by the primary, the quotient is called the *coefficient of coupling.* For example, when 10,000 lines of force are developed by the primary and 8,000 lines of force actually cut the secondary, the coefficient of coupling can be found as follows:

$$k = \frac{8{,}000}{10{,}000}$$

$$k = .8$$

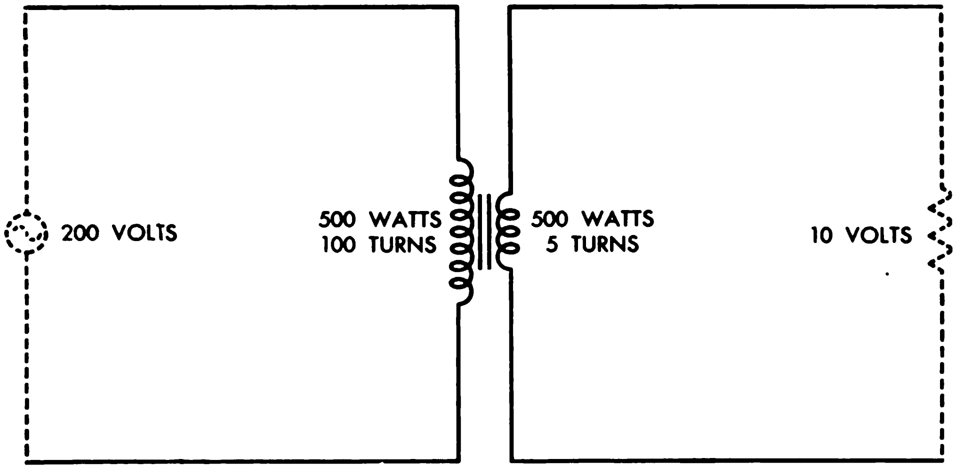
To find the true mutual inductance of a transformer, it is necessary to multiply the right-hand side of the previous equation for M by the coefficient of coupling, resulting in the formula,

$$M = k\sqrt{L_p \times L_s}$$

TURNS RATIO

If all of the lines of force from a transformer primary could cut all the turns of the secondary, and the copper and iron losses were negligible, the transformer would be 100 percent efficient. Assuming that a transformer is 100 percent efficient, and 500 watts of power is supplied to the primary, 500 watts of power must also be delivered to the secondary. An important transformer ratio can be established from this fact.

Suppose that the primary voltage of the transformer shown below is 200 volts and there are 100 turns of wire in the primary but only 5 turns in the secondary. Since the magnetic field from the 100-turn primary cuts only 1/20th as many turns in the secondary, the voltage induced in the secondary will be only 1/20th of the primary voltage or 10 volts. (The amount of induced voltage depends upon the number of turns.) Even though the secondary voltage is only 10 volts, the secondary power



Turns Ratio

must still be 500 watts. Therefore, the secondary current must be 50 amperes ($I = \frac{W}{E} = \frac{500}{10} = 50$ amperes). Thus, this transformer has stepped the voltage down from 200 to 10 volts and in so doing has stepped the current up from 2.5 to 50 amperes.

The relationship that exists between the primary and secondary voltages and the number of primary and secondary turns is usually expressed by the ratio,

$$\frac{E_p}{E_s} = \frac{N_p}{N_s}$$

where E_P is the primary voltage, E_S is the secondary voltage, N_p is the number of turns in the primary, and N_s is the number of turns in the secondary.

The relationship that exists between the number of primary and secondary turns and the primary currents and secondary currents is expressed by the ratio,

$$\frac{N_p}{N_s} = \frac{I_s}{I_p}$$

Note that the ratio of primary voltage to secondary voltage is equal to the ratio of primary to secondary turns, whereas the ratio of currents is equal to the *inverse* of the ratio of primary to secondary turns. If the primary has 10 turns and the secondary has 100 turns, the secondary voltage is *10 times* that of the primary, and the current in the secondary is *1/10th* of that in the primary.

A transformer that has more turns in the secondary than it has in the primary is called a *step-up* transformer. The secondary voltage of a step-up transformer is greater than its primary voltage. A *step-down* transformer, on the other hand, has less turns in the secondary than it has in its primary, and its secondary voltage is less than its primary voltage. Both step-up and step-down transformers are commonly used to increase or decrease values of AC voltages.

From the formulas relating current, voltage, and turns of transformers, you can see that the actual number of turns need not be known. The only necessary information about the turns is their *ratio*. The *turns ratio* of a transformer is defined as the ratio of the number of turns of the higher voltage winding to that of the lower voltage winding. As an example, the transformer that consists of 20 turns in one coil and 100 turns in the other has a 5:1 (5 to 1) turns ratio. If the primary has the smaller number of turns, it is a 5:1 step-up transformer. If the primary has the greater number of turns, it is a 5:1 step-down transformer. Therefore, two facts must be given to fully describe transformer ratios:

(1) the numerical ratio

(2) whether the transformer is step-up or step-down

The primary of an ideal transformer is connected to a 110-volt AC generator and the secondary is connected to a 5000-ohm resistor. If the primary has 60 turns and the secondary has 300 turns, what are the values of the secondary voltage, secondary current, and primary current?

Solution:

To calculate the secondary voltage:

$$\frac{E_p}{E_s} = \frac{N_p}{N_s}$$

$$\frac{110}{E_s} = \frac{60}{300} = \frac{1}{5}$$

$$E_s = 550 \text{ volts}$$

To determine the current flow in the secondary by Ohm's law:

$$I_s = \frac{E_s}{R_L}$$

$$I_s = \frac{550}{5500} = .11 \text{ ampere}$$

To determine the current in primary winding:

$$\frac{I_p}{I_s} = \frac{N_s}{N_p}$$

$$\frac{I_p}{.11} = \frac{300}{60} = 5$$

$$I_p = .55 \text{ amp}$$

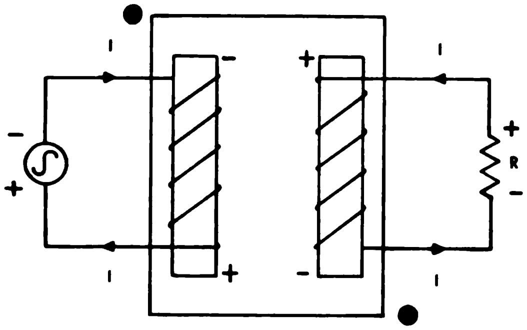
Voltage Phase Relationships

The voltage phase relationship between the primary circuit and the secondary circuit in an ideal untuned transformer (one having a resistive load) may be either 0° or 180°. This phase relationship depends upon two factors:

1. The manner in which the resistive load is connected to the secondary.

2. The manner in which the secondary is "wound" relative to the primary.

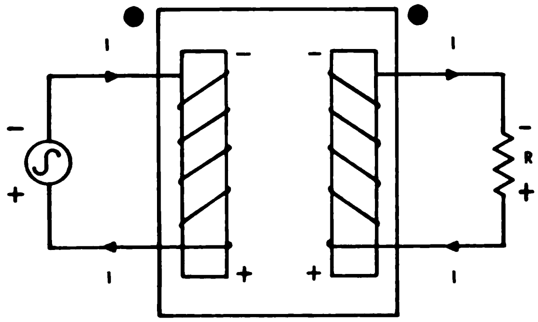
The figure below shows a typical transformer.



180° Out-of-Phase Transformer

The dots on the transformer schematic indicate that the top of the secondary is of opposite polarity to the top of the primary. This means that the unloaded secondary voltage is 180° out of phase with the primary voltage. Since the secondary is acting as a voltage source, current will flow *away* from the negative end of the secondary winding. As can be seen from the illustration, there would be a 180° phase shift in voltage from the generator to the load. If it were desirable to have the load voltage *in phase* with the generator voltage, it would only be necessary to reverse the secondary lead wires to the load resistor thereby reversing the direction of current flow through the load.

Another method by which to change the voltage phase between generator and load would be to rewind the secondary of the transformer. The transformer schematic would then be shown as in the figure below. (Note the change from the previous illustration in the position of the dots.)



In-Phase Transformer

It should be mentioned here that most transformers are wound so that the polarity across the primary winding is opposite to that across the secondary. If the dots are not shown on a transformer schematic, it is assumed that the phase shift, from *primary* winding to *secondary* winding is 180°.

IMPEDANCE RATIO

A transformer ratio that is very important in the study of communications equipment is the *impedance ratio*. It can be shown that the impedances of two circuits must be equal in order for the maximum amount of energy to be transferred from one circuit to the other. A transformer is often used to match two impedances that are not equal. Since the primary and the secondary of a transformer are coils, both have an impedance to alternating current. Actually, the square of the transformer-turns ratio is equal to the *impedance ratio* from primary to secondary. Expressed as a formula,

$$\left(\frac{N_p}{N_s}\right)^2 = \frac{Z_p}{Z_s}$$

where N_p is the number of turns in the primary, N_s is the number of turns in the secondary, Z_p is the impedance of the primary, and Z_s is the impedance of the secondary. For example, suppose you want to find the turns ratio of a transformer designed to match a 1250-ohm circuit (primary) to a 50-ohm circuit (secondary). Substituting in the above equation,

$$\left(\frac{N_p}{N_s}\right)^2 = \frac{1250}{50}$$

$$\left(\frac{N_p}{N_s}\right)^2 = 25$$

$$\frac{N_p}{N_s} = 5$$

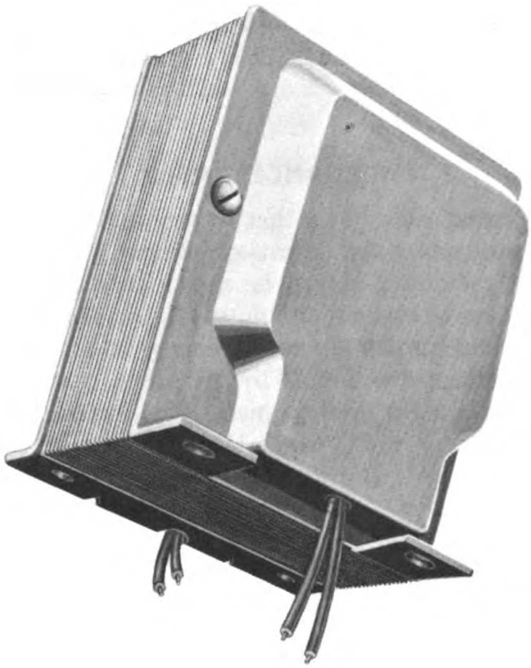
If a 1,000-turn primary is used in the 1250-ohm circuit, the secondary requires 200 turns since there must be five times as many turns in the primary as in the secondary.

TYPICAL TRANSFORMERS

There are three main types of transformers used in communications equipment. These are the *power transformer*, the *audio transformer*, and the *radio-frequency transformer*. Each of these types is designed for a special purpose.

Power Transformer

A power transformer is used to either step up or step down an AC voltage having a power frequency. Some power transformers have a



number of secondary windings each having a different number of turns. Some of the secondaries have fewer turns than the primary and, therefore, produce a lower voltage, and others have more turns and produce a higher voltage than that applied to the primary. A schematic of a common type of power transformer used in receivers is shown in the illustration at the right.

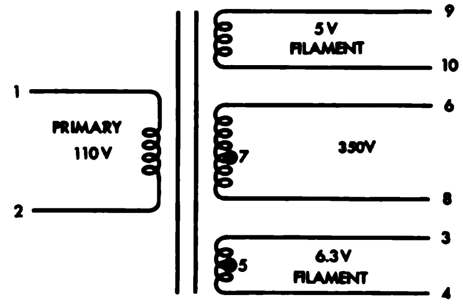
It is designed for a 110-volt primary voltage and has three secondary windings. A 6.3-volt secondary supplies current for heating the amplifier-tube heaters, a 5.0-volt winding supplies current for heating the rectifier-tube heater, and a 350-volt winding supplies the high AC voltage for the rectifier. (The rectifier produces a high DC voltage that is required by the amplifier tubes.)

A typical power transformer has color-coded leads so that it is easy to identify the primary and secondaries. The colored leads of the transformer in the schematic are shown in diagram B. The back leads are the primary; the red leads are the high-voltage secondary with a red-orange center tap; the yellow indicates the 5-volt secondary; and the green is the 6.3-volt winding with a green-yellow center tap. If the leads are not color-coded, they are numbered on the transformer and in schematics of the transformer.

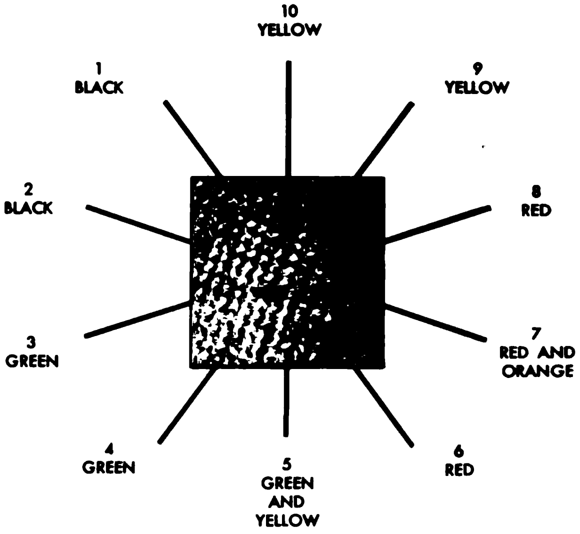
A power transformer has a laminated iron core similar to the core of the power choke discussed in chapter 13. This makes it possible to have a strong magnetic field from the primary to the secondary and, consequently, considerable power delivered to the secondary with a minimum power loss. The size of wire used in the different transformer windings depends primarily upon the amount of current that the particular winding must be able to carry. The insulation on the various windings depends upon the amount of voltage that will exist between the turns of the same winding, between the windings and the core, and between the different windings.

Audio Transformer

An audio transformer may be designed to merely transfer or *couple* an audio voltage or *audio signal* from one part of the equipment to



Power Transformer



Schematic and Color Coding of Power Transformer another, or it may be designed to step up or step down an audio voltage. For coupling, the audio transformer is designed with a 1:1 ratio — that is, there are the same number of turns in the primary as in the secondary. For step-up or step-down functions, the ratios may be as high as 5:1. Consequently, the audio transformer is usually used to transfer voltages of wide audio-frequency range, rather than voltages of a single frequency, as in the case of a power transformer.

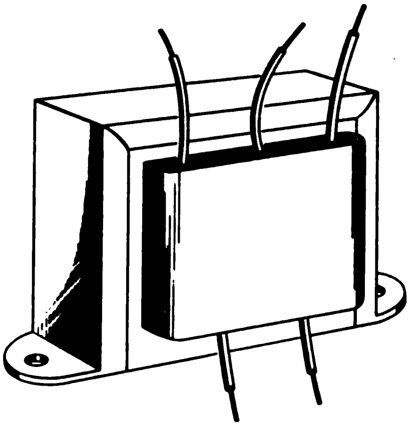
Audio transformers also have laminated iron cores and must be able to carry a limited amount of direct current in the primary windings without affecting AC audio frequency. However, the amount of iron in an audio-transformer core need not be as great as that in a power transformer, since the amount of power and, consequently, the number of lines of force developed in the core need not be as great. The size of the wire used in these transformers is quite small since the amount of audio current that flows in audio circuits is relatively small. The insulation requirements depend upon the amount of voltage developed across the turns in each particular case.

Radio-Frequency Transformer

Radio-frequency transformers are designed to couple RF energy from one circuit to another, to step up or step down RF voltages, or both. The number of turns in the primary and secondary of a particular transformer of this type depends upon the desired step-up or step-down ratio and the frequency of the applied voltage.

Radio-frequency transformers are quite small in size and usually are of the air-core type, but some may have powdered-iron cores. The amount of energy in an RF circuit is usually very small, and losses must be held to a minimum. Hysteresis losses and eddy-current losses are negligible in air cores and very small in powdered-iron cores. RF transformers may be *tuned* to one frequency or *untuned*.

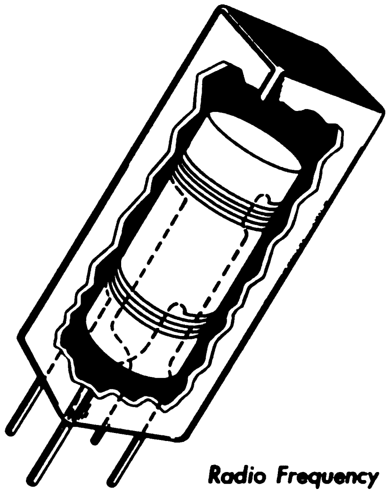
When radio-frequency transformers are used for transferring radio frequencies which are lower than the signal frequency (as in the case of a superheterodyne receiver, which is discussed on page 272), the device is known as an intermediate-frequency (IF) transformer. IF transformers may be tuned or untuned and they may have powdered-iron or air cores.



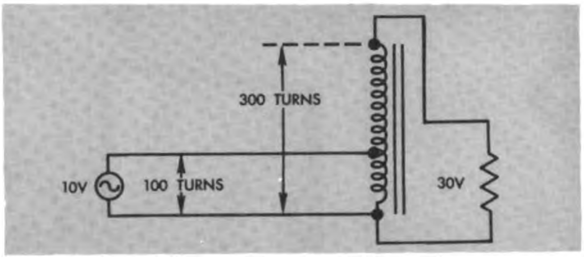
Audio Transformer

Autotransformer

A type of transformer that is commonly used in power circuits and is sometimes used in audio- and radio-frequency circuits is the *autotransformer*. An autotransformer circuit appears on the following page. The distinguishing feature of this type of transformer is that it has only one winding. In the illustration, there are 300 turns of wire on the iron-core. Each end of the winding is brought out to a terminal. A connection is made at a point 100 turns from one end of the winding. This is effectively the primary. If you wish to use this autotransformer to step up a voltage, apply the voltage to the 100 turns. The current through these turns sets up a magnetic field that cuts the other 200 turns, inducing the voltage across the remainder of the winding. The total voltage across the whole winding is equal to the sum of the applied voltage and the induced voltage. As shown in the illustration, if the applied voltage is 10 volts and the ratio is 3 to 1,



Radio Frequency Transformer



Autotransformer Circuit

the output voltage is 30 volts. Other taps may be used for different ratios. The main advantage of this type of transformer is that the primary acts as part of the secondary; and, therefore, less wire is required than in a standard transformer, which has a separate primary and secondary.

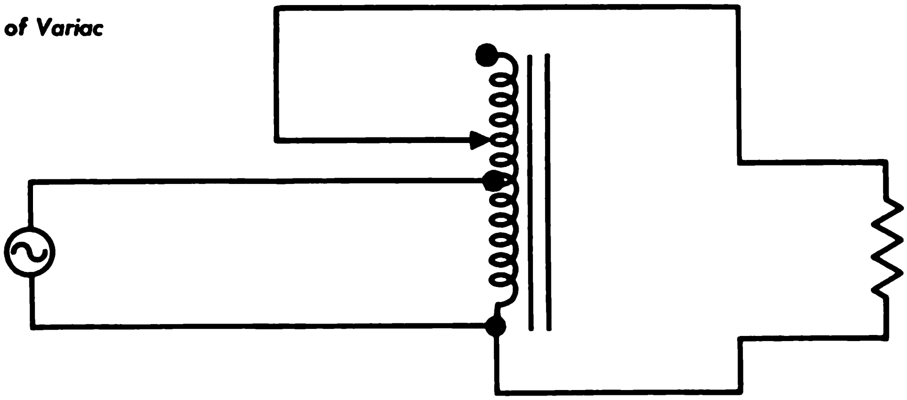
Another type of autotransformer is the *variac.* This name is a contraction of *variable* AC. The circuit of a variac is shown at the bottom of this page. One side of the output circuit is permanently connected to one end of the winding. The other side of the output is designed in a circular shape, so that the variable tap can be moved by a control knob that is located in the center. In the variac illustrated, the tap can be moved from one end of the winding to the other. This permits a setting of output voltage from zero to the maximum value and allows the variac to be used as a step-up or step-down transformer.

TESTING TRANSFORMERS

Simple checks can be made on transformers, similar to the checks that were discussed with respect to inductors in chapter 13. The most common check is for an *open transformer winding.* Since a winding is a continuous length of wire, it should have a definite amount of resistance. If it shows infinite resistance when measured with an ohmmeter (on the high scale), the winding must be open. Open windings are usually the result of excessive current. This causes the wire to melt and separate at a certain point or oxidize so much that it becomes thin and breaks.

A number of different types of *shorts* can develop in a transformer. The different windings can become shorted together. This trouble can be located by checking the resistance between the windings. The resistance between any two windings should be infinite. A winding can also become shorted to the transformer core. This type of trouble can be easily located since the resistance between any winding and the core should be infinite. One of the most difficult types of shorts to locate are shorts between a number of turns in the same winding. If the normal resistance of a winding is known, its resistance can be compared with this normal resistance value. For example, if the resistance should be 100 ohms but it measures only 50 ohms, approximately half of the turns are shorted together. On the other hand, if the normal resistance of a particular winding is small, it is very difficult to determine whether or not turns are shorted since the sensitivity of the ohmmeter is not great enough. Other useful checks will be encountered in the study of equipment that contains transformers, such as checking for the proper secondary voltages when power is being supplied to the transformer.

Circuit of Variac



Chapter 16

AC Series Circuits

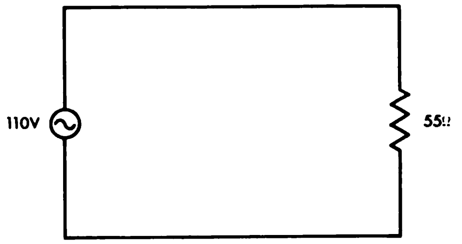
A series circuit is a circuit in which there is only one path for current to flow. The simplest series AC circuit is one in which a single resistor, inductor, or capacitor is connected to the power source. In a circuit of this type, there is only one value of current — the same current flows through the power source and the component. If the component is an inductor, the current lags the applied voltage by 90° and the opposition to the current is called inductive reactance. A circuit of this type dissipates very little power since the inductor stores energy in its magnetic field while the current through it is increasing, and delivers the energy back to the circuit while the current is decreasing. Since inductors consist of a number of turns of wire, all inductors have a small amount of resistance in addition to their inductance; therefore, some energy is dissipated in the resistive component. If the component is a capacitor, the current leads the applied voltage by 90°, and in this case, the opposition to the current is called capacitive reactance. This type of circuit also dissipates very little power since the capacitor stores energy in its electric field while the applied voltage is increasing, and delivers the energy back to the circuit while the voltage is decreasing.

Another type of series AC circuit is one that consists of an inductor and a capacitor connected in such a manner that the current that flows through the inductor must also flow through the capacitor. Inductive reactance cancels the effect of the opposition due to the capacitive reactance and vice versa. If the reactances are equal, there is complete cancellation. This type of circuit is called a resonant circuit and is treated fully in chapter 17.

This chapter deals with AC circuits, some of which consist only of resistance and others, of combinations of resistance and reactance. Relatively simple methods for solving series AC problems are discussed in regard to each type of circuit.

RESISTIVE CIRCUIT

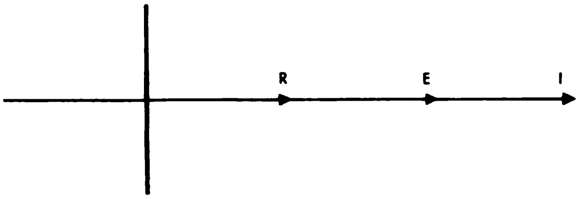
A series circuit that contains only resistance and a power source is called a resistive circuit. Since a resistor does not store energy as do inductors and capacitors, the current through a resistor is always in phase with the voltage across the resistor. Therefore, a resistive AC circuit is treated in the same manner as a DC circuit. For example, suppose that a 55-ohm resistor is connected across a 110-volt AC line as shown below. You can find the current flow in this circuit by Ohm's law as follows:



$$I = \frac{E}{R}$$
$$= \frac{110}{55}$$
$$= 2 \text{ amperes}$$

Since 110 volts is the effective value of voltage, 2 amperes is also the effective value of current. You can find the amount of power dissipated by the resistor by any one of the power formulas that was used in conjunction with DC circuits, provided that effective values are used. For example,

$$P = E \times I$$
$$= 110 \times 2$$
$$= 220 \text{ watts}$$



Phase Relationship in an AC Resistor Circuit

If a resistive AC circuit is made up of two or more resistors connected in series, the total resistance is the sum of the individual resistances. This sum is then used in finding the current flow.

Since voltage, current, and resistance have magnitude and direction, these quantities may be represented by vectors for a series circuit. All vectors are drawn with reference to the current vector which is drawn in the horizontal position along the X axis, as shown in the illustration to the left below.

Since the current in the circuit reaches maximum and minimum values at the same time as the applied voltage and the voltage across the resistor, they are all in phase. Each vector shown may be represented in polar form; that is, $E = 110/0°$ $I = 2/0°$, and $R = 55/0°$. This merely means that the voltage vector is drawn 110 voltage units along the horizontal axis at an angle of 0°, that the current vector is drawn 2 current units along the horizontal axis at 0°, and that the resistor is drawn 55 resistance units along the horizontal axis at 0°. Zero degrees is always represented along the horizontal axis extending to the right. A discussion of vectors and their representation in polar form is given on pages 533 to 536.

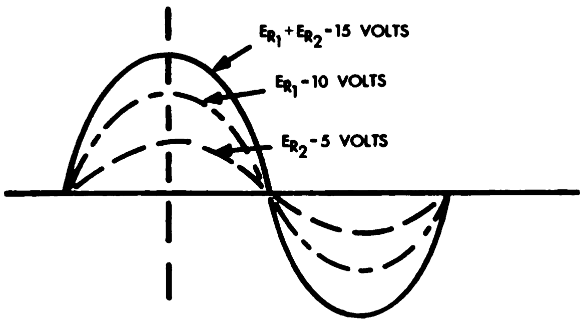
RL CIRCUITS

When an inductor and a resistor are connected in series to a source of alternating current the same current flows through both components. This type of circuit is called a series RL circuit. However, the voltage developed across the inductor, because of this current flow leads the current by 90°, while the voltage across the resistor is in phase with the current. Kirchoff's voltage law for direct-current circuits states that the sum of the voltage drops around a circuit equals the applied voltage. This is true for a resistive AC circuit, but not for a circuit made up of a resistance and a reactance (inductive or capacitive reactance) as shown in the illustration on the next page.

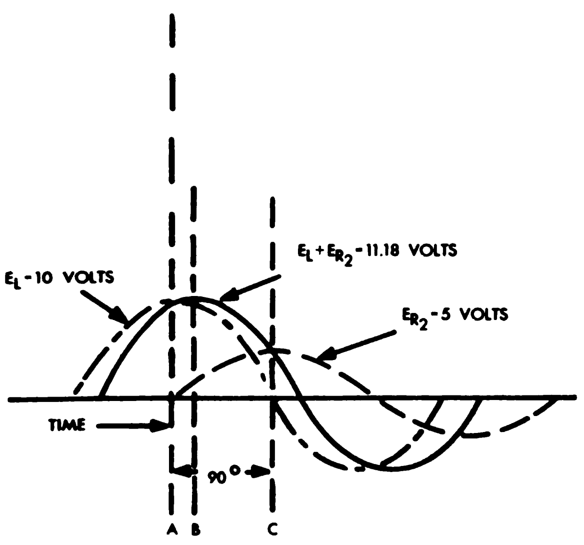
The sine-wave voltages across two resistors in a series circuit are shown in diagram A. Since the two voltages are in phase, they can be added directly. They reach their maximum values together at time A and, therefore, the sum of their peak values is:

$$E_{R_1}+E_{R_2}=10+5$$
$$=15 \text{ volts}$$

In diagram B, two voltages equal to those in the upper drawing, but 90° out of phase with each other, are added to compare the total voltages. At time A, $E_L=10$ volts and $E_R=0$ volts. The sum at this time is $10+0=10$ volts. At time B, E_L has decreased to 8.94 volts and E_R has increased to 2.24 volts. Adding these together, the instantaneous voltage across both components at time B is found to be 11.18 volts. This is the peak value of the voltage sums. At time C, E_L has decreased to 0 volts and E_R has increased to its peak of 5 volts. In summary, the resultant of 5 volts (peak) and 10 volts (peak) is 15 volts (peak) when the two voltages are in phase. When those same two voltages are 90° out of phase, the peak value of their sums is only 11.18 volts.



RESISTIVE-CIRCUIT WAVEFORMS



RL CIRCUIT WAVEFORMS

Effect of Phase Angle on Total Voltage

Whenever the vector sum of the voltages across all the components in a series circuit *leads the current* flowing through the circuit, the circuit is said to be *inductive*. If the amount of resistance is large in comparison to the amount of inductive resistance the phase angle will be small (near 0°). If the amount of inductive reactance is large in comparison to the amount of resistance, the phase angle will be large (near 90°).

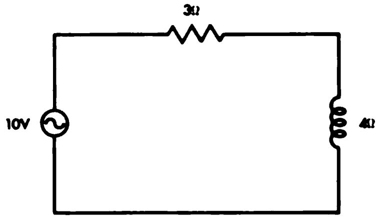
In dealing with series AC circuits, it is often necessary to solve for unknown voltages, impedances, current, phase angle, power, and power factor. These solutions will be discussed in reference to voltage and impedance. Power factor is discussed separately.

Solving for Impedance

It is an easy matter to determine the amount of resistance in the case of the series resistive circuit. You merely add the various resistances in the circuit. However, in a series AC circuit containing resistance and inductance, finding the total resistance to current flow is not so easy. The total opposition to current flow is now found by adding vectorially the resistance and inductive reactance contained in the circuit. This combined opposition to current flow is called impedance (Z).

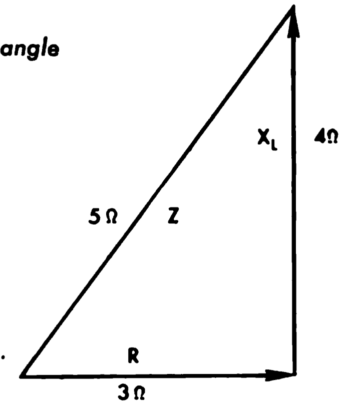
The total impedance of a resistor and an inductor can be found by vector addition. This is called a graphic solution and uses the impedance triangle as explained below.

GRAPHIC SOLUTION. Assume that a certain series RL circuit, as shown in the illustration below, consists of a power source generating 10 volts, a 3-ohm resistor and a coil having 4 ohms of inductive reactance. Suppose that you want to find the impedance of the circuit using the graphic method.



A Series RL Circuit

The Impedance Triangle



Draw a horizontal line to represent resistance, and label this line R. The length of this line must be determined by a convenient scale, say 1 ohm per inch. Since inductive reactance causes a voltage drop that leads the current, draw a vertical line upward from the right-hand end of line R and label it X_L. The same scale must be used in determining the length of this line. Now draw a line connecting the left end of the line R with the top of line X_L. Label this line Z. A triangle of this type is known as an impedance triangle and is shown in the illustration. When the length of line Z is converted to ohms by the scale used in determining the lengths of lines R and X_L, you will find that the total impedance is 5 ohms. The angle between line R and line Z is the phase angle (θ) that this combination of resistance and inductive reactance will cause, when current flows through the circuit. This angle should measure about 53°.

As a navigator, you have at your disposal an instrument — a DR computer — that will solve this graphic solution for you very easily. On the wind vector side of the computer, place the grommet over the 50* on the grid. Place N under the true index, and plot the resistance 3 units out from the grommet, and the inductive reactance 4 units up from the end of the resistance vector as shown in the illustration to the right. Now join the ends of the vectors, and the impedance triangle is complete.

To find the magnitude of Z and its phase angle, rotate the compass rose clockwise until the Z line is horizontal as shown in the illustration at top of page 119. The magnitude of Z is found by counting to the end of Z (5 units). By the scale used, Z is equal to 5 ohms. The phase angle θ can be determined by counting counterclockwise the number of degrees from N to the true index. In this case, $\theta = 53°$.

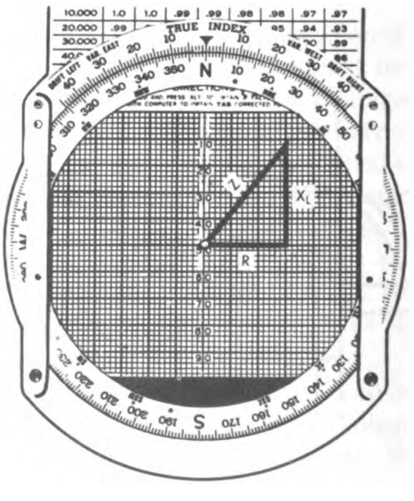
SOLUTION BY THE RIGHT-TRIANGLE EQUATION. Another method that can be used to find the impedance in this example uses a relationship that exists in a right triangle. However, before studying this method, you should first review the characteristics of a right triangle.

When using this method, sketch a right triangle and label the legs and hypotenuse according to the impedances that they represent. The horizontal leg is always labeled R for resistance, and the vertical leg is labeled X_L for inductive reactance. In this case the resistance is 3 ohms and the inductive reactance is 4 ohms. The hypotenuse represents the total impedance and is labeled Z. To find Z,

$$\begin{aligned} Z &= \sqrt{R^2+X_L^2} \\ &= \sqrt{3^2+4^2} \\ &= \sqrt{9+16} \\ &= \sqrt{25} \\ &= 5 \text{ ohms} \end{aligned}$$

Solving for Current

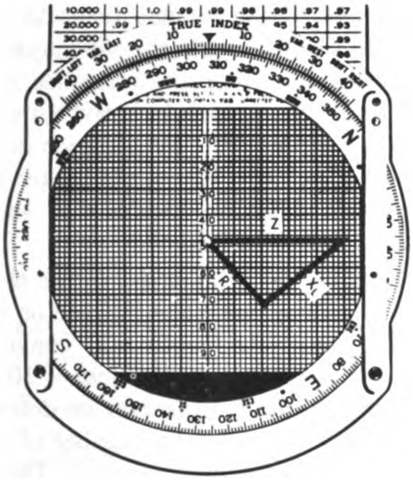
In this series resistive circuit, you merely used Ohm's law to solve for I. You can do the same in a series RL circuit, but be sure to use the



Drawing the Impedance Triangle Using the DR Computer

---

*For other type computers, place the grommet on any convenient line.



Obtaining the Magnitude of Z from the DR Computer

total impedance Z that is offered to the flow of current, not just the resistance of the circuit. Ohm's law for AC circuits containing reactance is E = IZ. Solving this formula for current, you find I = E/Z. In the RL circuit problem, the total impedance of the circuit is 5 ohms. Using this value in Ohm's law,

$$I = E/Z$$

$$= \frac{10}{5}$$

$$= 2 \text{ amps}$$

Notice that, although the solution gives the magnitude of the current, it does not give the phase angle. You can find the phase angle by writing the voltage and impedance in polar form in the formula for Ohm's law and solving for the current I as follows:

$$I = \frac{100\underline{/0^\circ}}{50\underline{/53^\circ}}$$

$$= 2\underline{/-53^\circ}$$

This solution shows that the current of 2 amps lags the applied voltage by 53°. A discussion of the use of the polar form in vector division is on page 535.

Solving for Voltage

Once the total impedance and current of the series RL circuit is obtained, it is an easy matter to find the voltages existing across the resistor and across the inductor. The voltage across the resistor is:

$$E_R = IR$$

$$= 2 \times 3$$

$$= 6 \text{ volts}$$

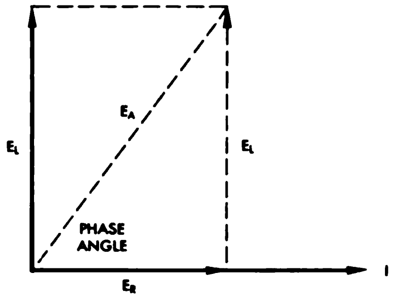
The voltage across the inductor is:

$$E_L = IX_L$$

$$= 2 \times 4$$

$$= 8 \text{ volts}$$

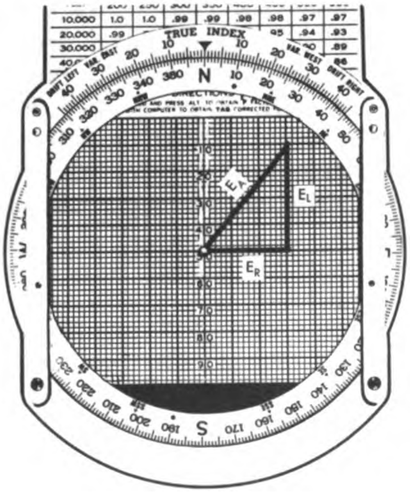
GRAPHIC SOLUTION. Notice that at first glance this solution seems to be wrong since $E_R + E_L = 14$ volts which exceeds the applied voltage of 10 volts. However, as pointed out earlier, these voltage drops do not occur at exactly the same time; therefore, they must be added vectorially, as shown in the illustration below.



Addition of Voltage Vectors in an RL Series Circuit

Since the current is the same through both components, draw a vector representing the current and use this vector as a reference as shown. The length of this current vector has no relationship to the length of the voltage vectors since they are not measured in the same units. Because the voltage across and the current through the resistor are in phase, the vector representing the voltage (E_R) across the resistor is drawn along the current vector. This vector must be drawn to scale in respect to the amount of voltage that it represents. If you choose a scale of 2 volts per inch, for example, E_R should be three inches long. (Any other convenient scale can be used.) Since the voltage across the inductor leads the current by 90°, the angle between vector I and the vector representing the voltage (E_L)

across the inductor must be 90°. A leading vector is drawn counterclockwise from the reference vector, which places vector E_L vertically upward from the center of the vector diagram. Use a protractor to construct the 90° angle accurately. The length of the vector E_L must be determined by the same scale chosen for vector E_R. If you use 2 volts per inch, then the length of E_L should be 4 inches.



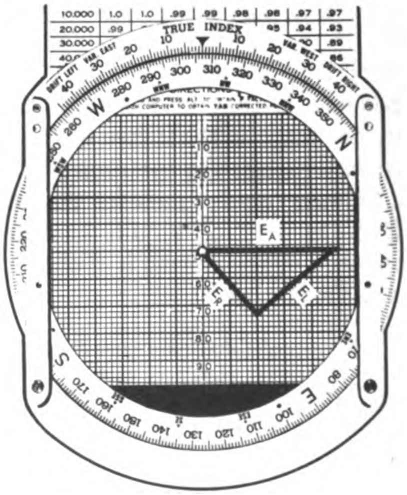
Addition of Voltage Vectors on the DR Computer

Two vectors are added together by placing the tail of one next to the head of the other, without changing the direction of either in respect to the reference vector. This can be done by moving vector E_L. The vector sum (E_A or applied voltage) is the distance from the tail of E_R to the head of dotted vector E_L. If you use a scale of 2 volts per inch, this vector sum should measure 5 inches in length. This length must be converted to volts by the same scale that was used in constructing the individual voltage vectors. If you are using a scale of 2 volts per inch, E_A equals 10 volts. Even though a different scale is used, the result should be the same.

The phase angle between the applied voltage and the current through the circuit is the angular distance from vector I to the resultant vector E_A. When you measure this angle with a protractor, it should be approximately 53°.

This problem can be solved on the DR computer in the same manner as the impedance triangle. Place N under the true index on the wind side and the grommet over 50 on the grid. Choose a convenient scale (1 major square equals 2 volts) and draw the reference line current out from the grommet. Draw on this line E_R (6 volts), represented by 3 major divisions. From the end of the E_R vector draw the E_L vector upward 4 major divisions as shown in the illustration to the left.

To determine the magnitude, rotate the compass rose clockwise until the E_A vector is horizontal. Determine its length and convert this length to voltage (5 major divisions = 10 volts). Notice that the phase angle can be determined at this time by counting the number of degrees the compass rose has been rotated. The phase angle is 53° as shown in the illustration below.



Computing Magnitude of Phase Angle from DR Computer

SOLUTION BY RIGHT TRIANGLE EQUATION. Since voltage vectors E_R, E_L, and E_A in the preceding series RL circuit form a right triangle, the applied voltage (E_A) can be found by the right-triangle equations as follows:

$$\begin{aligned} E_A &= \sqrt{E_L{}^2+E_R{}^2} \\ &= \sqrt{8^2+6^2} \\ &= \sqrt{64+36} \\ &= \sqrt{100} \\ &= 10 \text{ volts} \end{aligned}$$

Power Factor

As you have learned, an inductor stores energy during the part of the cycle in which current is increasing and delivers most of this energy back to the circuit when the current is decreasing. Consequently, to find the power dissipated in the RL circuit, it is necessary to consider only the resistance. Solving for power we get

$$\begin{aligned} P &= I^2 R \\ &= (2)^2 \times 3 \\ &= 4 \times 3 \\ &= 12 \text{ watts} \end{aligned}$$

This value is called the *true power* dissipated in the circuit (resistor) in the form of heat. Power formulas $P = EI$ and $P = \frac{E^2}{R}$ can be used only when the voltage across the resistor (E_R) is substituted for E.

If the square of the current through an inductance-resistance circuit is multiplied by the total impedance, the product is called the *apparent power.*

$$AP = I^2 Z$$

This value of power includes the power that is being stored in the reactance component and is not the actual power dissipated in the circuit. The ratio of the value of true power to the value of apparent power is called the *power factor* of the circuit. In formula form,

$$PF = \frac{TP}{AP}$$

A simplified version of this formula can be derived by substituting the values of true power and apparent power in the above formula.

$$PF = \frac{TP}{AP}$$

$$PF = \frac{I^2 R}{I^2 Z}$$

Since the value of I^2 in the numerator and denominator is the same,

$$PF = \frac{R}{Z}$$

where R is the resistance in the circuit and Z is the total impedance of the circuit.

Note that the value $\frac{R}{Z}$ is equal to the cosine of the phase angle of the vector triangle. Therefore, $PF = \cos \theta$.

In a purely resistive circuit, where the impedance and the resistance are the same, the power factor is one or unity. In a purely inductive circuit, the resistance is zero and, therefore, the power factor is zero. In the simple resistance-inductance circuit discussed previously, the resistance is 3 ohms and the inductive reactance is 4 ohms. In this circuit, the power factor is found in either of the following ways:

$$\begin{array}{r|r} PF = \frac{R}{Z} & PF = \cos \theta \\ = \frac{3}{5} & = \cos 53° \\ = .6 & = .6 \end{array}$$

When the power factor is one, the apparent power is equal to the true power. Since only true power does useful work, it is desirable to have the power factor as near 1 as possible.

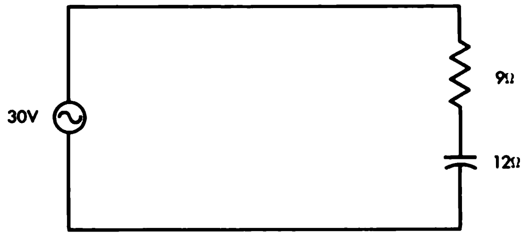
RC CIRCUITS

When a capacitor and a resistor are connected in series with a power source, the same current flows through both components as it does when an inductor and a resistor are connected in series. Although electrons do not actually pass through a capacitor, the effect is the same because of the charge and discharge of the capacitor. The voltage that is developed across the capacitor lags the current by 90° (assuming a pure capacitance) while the voltage across the resistor is in phase with the current. The net result in an RC circuit is that the applied voltage lags the current by an angle between 0° and 90°, depending upon the relative amounts of capacitive reactance and resistance. If the capacitive reactance is much greater than the resistance, the phase angle is near 90°, but if the resistance is much larger than the capacitive reactance, the phase angle is near 0°. Regardless of the relative amounts of each, the circuit is said to be capacitive since the applied voltage lags the current.

As in the case of RL circuits, it is often necessary to solve for voltage, impedance, current, phase angle, power, and power factor in RC circuits. These solutions are found in much the same manner as they were found in RL circuits.

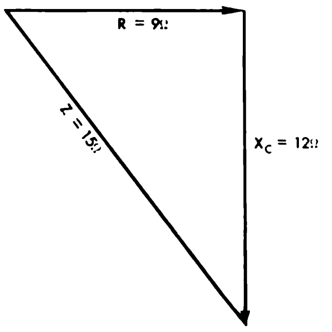
Solving for Impedance

Assume that a series circuit, as shown in the illustration at the top of page 122, is made up of a 30-volt power source, a capacitor whose reactance at the applied frequency is 12 ohms, and a 9-ohm resistor. Let us find how much current is flowing in the circuit.



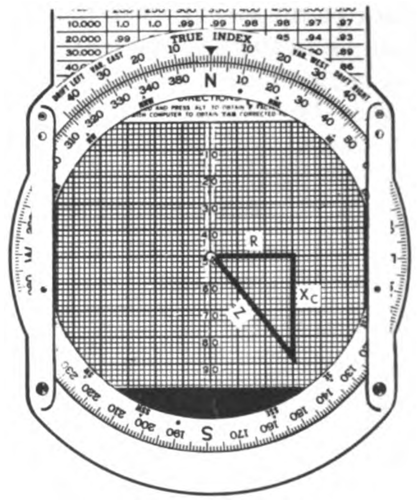
A Series RC Circuit

GRAPHIC SOLUTION. As in the series RL circuit, first find the total impedance offered to the flow of current. This may be done by constructing an impedance triangle as shown in the illustration below. The resistance is again used as our reference line and the vector representing it is drawn 3 units long with each unit equal to 3 ohms. Since a capacitive circuit causes the current to lead the voltage, the X_C vector is drawn downward from the end of the resistance vector and 4 units long. The amount of impedance offered to the flow of current is then represented by the length of the line joining the ends of the resistance and capacitance reactance vector. This is equal to 5 units or 15 ohms.



Impedance Triangle for RC Circuit

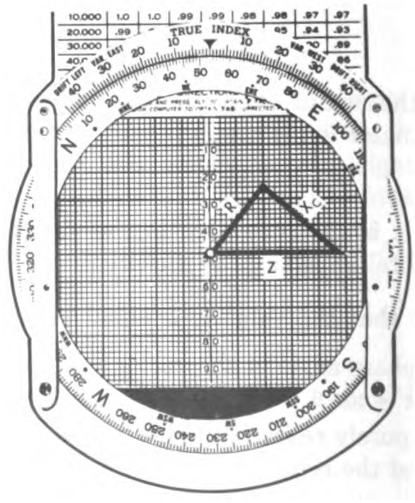
This impedance problem can also be solved on the DR computer. Place N of the compass rose under the true index with the grommet over the 50 on the grid. Choose a convenient scale (each major division equals 3 ohms) and plot the resistance out horizontally 3 major divisions. From the end of the resistance vector plot the X_C vector downward 4 divisions as shown in the next illustration.



The RC Impedance Triangle on the DR Computer

To determine the magnitude of the impedance, rotate the compass rose counterclockwise until the impedance line is horizontal as shown in the illustration below. The amount of impedance in the circuit can then be determined by multiplying the number of division lines by the scale used. In this case 5×3=15 Ω. The angle between R and Z can be found by reading under the true index, 53°.

SOLUTION BY RIGHT TRIANGLE EQUATION. You can solve for Z by applying the right-triangle equation to the impedance triangle.



Determining Magnitude of Z on the DR Computer

$$\begin{aligned} Z &= \sqrt{R^2+X_C^2} \\ &= \sqrt{(9)^2+(12)^2} \\ &= \sqrt{81+144=225} \\ &= 15 \text{ ohms} \end{aligned}$$

Solving for Current

Now that you know the total impedance, you can find the current flowing in the circuit in the same manner as in the series RL circuit.

$$\begin{aligned} I &= E/Z \\ &= \frac{30}{15} \\ &= 2 \text{ amps} \end{aligned}$$

The angle by which the current leads the applied voltage can be found by writing the voltage and impedance in polar form and solving for current. Since the angle between our reference vector and the impedance vector is measured clockwise, the polar form of Z is $15\underline{/-53^\circ}$. The current vector may be found as follows:

$$\begin{aligned} I &= \frac{E}{Z} \\ &= \frac{30\underline{/0^\circ}}{15\underline{/-53^\circ}} \\ &= 2\underline{/53^\circ} \end{aligned}$$

This vector shows that the current leads the voltage in a capacitive circuit.

Solving for Voltage and Phase Angle

Once the current that flows in the circuit is known, you can determine the voltage drops existing across the resistor and capacitor. The voltage drop across the resistor is

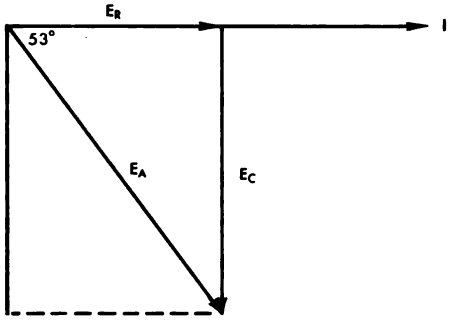
$$\begin{aligned} E_R &= IR \\ &= 2 \times 9 \\ &= 18 \text{ volts} \end{aligned}$$

The voltage across the capacitor is

$$\begin{aligned} E_C &= I\,X_C \\ &= 2 \times 12 \\ &= 24 \text{ volts} \end{aligned}$$

Notice that once again it appears that the sum of the voltage drops exceeds the applied voltage; however, remember that these voltage drops do not occur simultaneously and that they must be added vectorially.

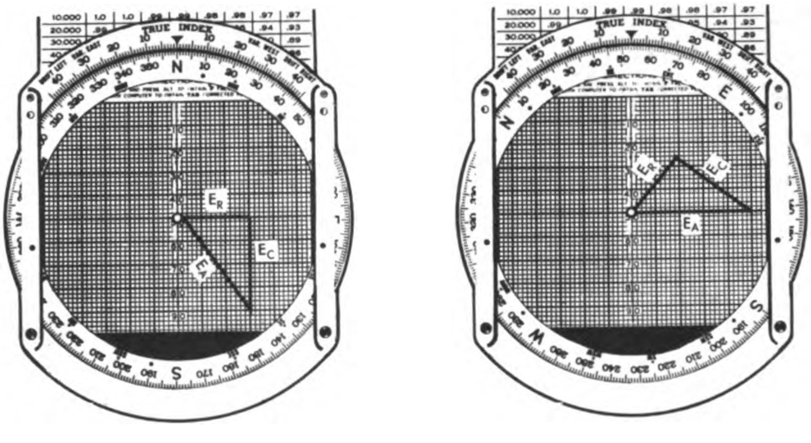
GRAPHIC SOLUTION. To add the voltage drops graphically, draw a current vector as a reference. Label this vector I, as in the following illustration. Remember that the length of this current vector does not enter into the solution, but it should be drawn horizontally and to the right from the center of the vector diagram. Since the voltage across the resistor is in phase with the current, draw a vector representing this voltage (18 volts) to scale, along the current vector. Label this voltage vector E_R. Now draw a vector representing the voltage across the capacitor and label this vector E_C. Since the voltage across the capacitor lags the current by 90°, draw this vector lagging vector I by 90° and equal to 24 volts. Complete the parallelogram or rectangle by drawing dotted lines parallel to vectors E_R and E_C. Draw the diagonal from the center of the vector diagram to the opposite corner of the rectangle. Label this diagonal E_A. Measure the length of this resultant vector and convert its length to voltage, using the same scale that you used for E_R and E_C. E_A is approximately 30 volts.



Vector Solution, RC Circuit

The phase angle is the angle between the current vector and the applied voltage vector. When measured with a protractor, this angle should be approximately 53°. Actually the angle is negative, indicating that the applied voltage lags the current by 53°.

The DR computer can be used to find the applied voltage in the following manner. On the wind side of the computer set N under the true index and the grommet over 50 on the grid. Choose a convenient scale (one large square equals 6 volts) and draw the voltage across the resistor horizontally from the grommet. The voltage across the capacitor lags the current by 90° and is therefore drawn downward from the end of the E_R vector, as shown in the illustration on page 124. The magnitude of E_R is found by rotating the compass rose counterclockwise until the E_A vector is horizontal and then converting its length into voltage. The phase angle can also be found at this time by reading the number of degrees under the



The Capacitive Voltage Vector on the DR Computer

true index. In the illustration given, E_A is equal to 30 volts and the current leads E_A by 53°.

Solution by Right Triangle Equations

The applied voltage can also be found by using the right triangle equation. Since $E_C = 24$ volts and $E_R = 18$ volts, draw a right triangle with these values as legs of a right triangle. Then,

$$\begin{aligned} E_A &= \sqrt{E_R^2 + E_C^2} \\ &= \sqrt{(18)^2 + (24)^2} \\ &= \sqrt{324 + 576} = \sqrt{900} \\ &= 30 \end{aligned}$$

Power Factor and Power Dissipated

You can find the power factor of this circuit by using either of the two formulas:

$$PF = \frac{R}{Z} \qquad\Big|\qquad PF = \cos\theta$$

$$PF = \frac{9}{15} \qquad\Big|\qquad PF = \cos 53°$$

$$= .6 \qquad\Big|\qquad = .6$$

The power dissipated in this circuit, or the true power is equal to the square of the current times the resistance.

$$\begin{aligned} P &= I^2 R \\ &= (2)^2 \times 9 \\ &= 4 \times 9 \\ &= 36 \text{ watts} \end{aligned}$$

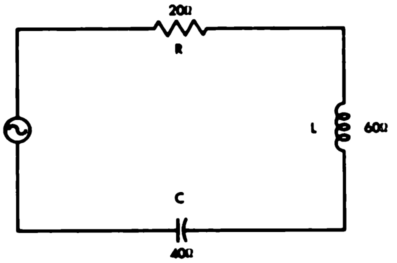
RCL CIRCUITS

As in the case of the other series AC circuits discussed, the same current flows through all three components when a capacitor, an inductor, and a resistor are connected in series (series RCL circuit). Here again, the electrons do not actually flow through the capacitor but the effect is the same. In this type of circuit, there are three individual voltages that are out of phase with one another. The voltage across the inductor leads the current by 90°, the voltage across the capacitor lags the current by 90°, and the voltage across the resistor is in phase with the current. Since one reactive voltage leads the current by 90° and the other lags the current by 90°, they are 180° out of phase with each other and are therefore in direct opposition. If these two reactive voltages were equal, the circuit would act as a resistive circuit with no reactance at all. Such a circuit is called a resonant circuit, several of which are discussed in chapter 18, Resonant Circuits.

Solving for various unknown values in series circuits made up of all three types of impedance can be accomplished by the methods discussed previously. Slightly different procedures are applied however, when two voltages are 180° out of phase and two impedances have opposing effects in the circuit.

Solving for Voltage

Assume that a power source is connected in series with a 20-ohm resistor, an inductor whose reactance at the applied frequency is 60 ohms, and a capacitor whose reactance at the same frequency is 40 ohms. Also assume that the current flowing through this circuit is .5 ampere. Find the applied voltage, the phase angle, the

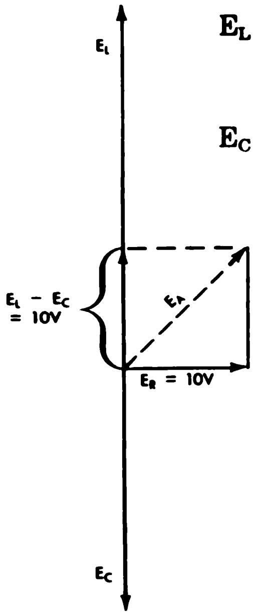


The RLC Series Circuit

power dissipated in the circuit, and the power factor.

GRAPHIC SOLUTION. The voltages across the individual components are:

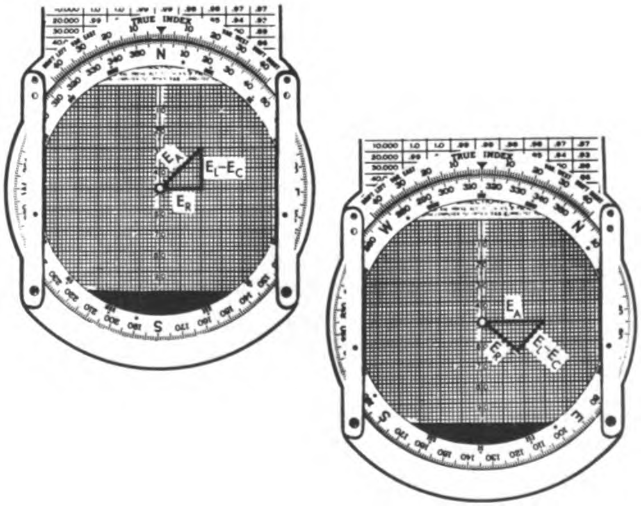
$$
\begin{aligned}
E_R &= IR \\
&= .5 \times 20 \\
&= 10 \text{ volts} \\
E_L &= IX_L \\
&= .5 \times 60 \\
&= 30 \text{ volts} \\
E_C &= IX_C \\
&= .5 \times 40 \\
&= 20 \text{ volts}
\end{aligned}
$$



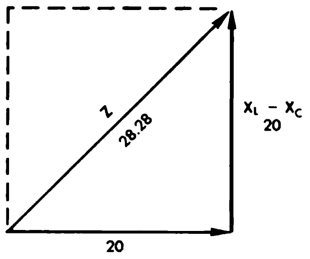
Vector Diagram of Series RLC Circuit

Construct a vector diagram of these individual voltages as shown in the illustration above, using the current as a reference vector. Remember that the voltage vectors must be drawn to scale. Using the selected scale, draw the vector representing the voltage across the resistor (E_R) to the right, the vector representing voltage (E_L) across the inductor leading the current vector by 90°, and the vector representing the voltage (E_C) across the capacitor lagging the current vector by 90°. Now find the resultant of the two opposing vectors by subtracting the length of the shorter from the length of the longer and giving the difference vector ($E_L - E_C$) given the direction of the longer. Add this difference vector ($E_L - E_C$) to vector E_R by the parallelogram method. Draw dotted lines parallel to vectors ($E_L - E_C$) and vector E_R, and draw the diagonal from the center of the vector diagram to the opposite corner of the parallelogram or rectangle. Convert the length of this vector to voltage, using the scale that you used to construct the voltage vectors, and you will find that the applied voltage E_A is approximately 14 volts. Since there was more inductive reactance in the circuit than capacitive reactance, the circuit is said to be inductive. In inductive circuits the applied voltage leads the current. In the illustration, the phase angle is 45°; therefore, the applied voltage leads the current by 45°.

This problem is solved on the DR computer in the same manner as the other problems. On the wind side, set N under the true index and the grommet over 50 on the grid. Choose a convenient scale (1 large block equals 5) and draw the voltage across the resistor horizontally from the grommet. Take the difference between the two reactances ($X_L - X_C$), and since X_L is the larger, plot the difference upward from the end of the E_R vector as shown in the illustration below. Rotate the compass rose clockwise until the E_A vector is horizontal and convert its length into volts. The phase angle is obtained by counting the number of degrees the compass rose has been rotated. The results are: $E_A = 14$ volts, the phase angle $= 45°$.



The RLC Voltage Vector on the DR Computer



RIGHT TRIANGLE SOLUTION. The right-triangle equation for finding the applied voltage across an RCL circuit when E_L is greater than E_C is

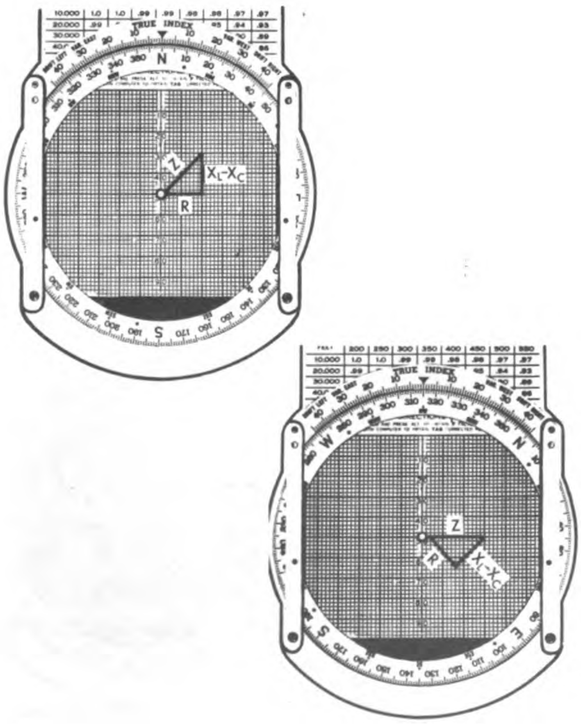
$$E_A = \sqrt{E_R^2 + (E_L - E_C)^2}$$

If the voltage across the capacitor is larger than the voltage across the inductor, the second term under the radical is $(E_C - E_L)^2$. Applying the above equation to the preceding RCL problem,

$$\begin{aligned} E_A &= \sqrt{10^2 + (30-20)^2} \\ &= \sqrt{10^2 + 10^2} \\ &= \sqrt{200} \\ &= 14.14 \text{ volts} \end{aligned}$$

Solving for Impedance

Assuming that the original problem was to find the total impedance, you can construct an impedance triangle, following a procedure similar to that for the RL circuit or the RC circuit.



The RLC Impedance Vector on the DR Computer

GRAPHIC SOLUTION. When constructing an impedance triangle for an RCL circuit, you must first find the total effective reactance (the difference between the inductive reactance and capacitive reactance). In this problem, this difference is an inductive reactance of 20 ohms. In other words, if the 60-ohm inductive reactance and the 40-ohm capacitive reactance are removed from the circuit, and a 20-ohm inductive reactance is placed in the circuit, the total current and the phase angle will remain unchanged. The impedance triangle can then be constructed with leg R and leg X $(X_L - X_C)$ drawn to scale as illustrated to the left. Impedance Z is the hypotenuse and its value can be determined from the scale that you have used. When Z is determined by this method, it is about 28 ohms.

The impedance triangle of an RCL circuit is constructed on the DR computer in the same manner as for the RC and RL circuits. On the wind vector side, place N under the true index and the grommet over 50 on the grid. Draw horizontally, to scale, the resistance in the circuit, 20 ohms. Obtain the difference between the inductive and capacitive reactances in the circuit. If the larger is inductive, as it is in this problem, plot the difference upward as shown in the diagram. If the larger is capacitive, plot it downward. To find the magnitude of Z, rotate the compass case clockwise until the impedance line is horizontal and convert its length to ohms as shown in the illustration to the left. The phase angle is found from the computer to be 45°.

RIGHT TRIANGLE SOLUTION. The right-triangle equation for finding the total impedance in an RCL circuit when X_L is greater than X_C is

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

If the capacitive reactance is greater than the inductive reactance, the second term under the radical is $(X_C - X_L)^2$. Substituting the values in the preceding problem in these equations,

$$\begin{aligned} Z &= \sqrt{20^2 + (60-40)^2} \\ &= \sqrt{400 + 400} \\ &= \sqrt{800} \\ &= 28.28 \text{ ohms} \end{aligned}$$

Chapter 17

Parallel AC Circuits

A parallel circuit is a circuit which consists of two or more paths for current to follow. The voltage drop across each component in a parallel circuit is equal to the applied voltage; therefore, all the voltage drops are equal. A parallel AC circuit may consist of a power source and two or more resistors, two or more inductors, or two or more capacitors. A parallel AC circuit may also consist of various combinations of resistors, inductors, and capacitors. Circuits that contain equal amounts of capacitive reactance and inductive reactance are discussed in chapter 18, Resonance.

Parallel resistive AC circuits are treated in the same manner as DC circuits. The total resistance is equal to the reciprocal of the reciprocals and the total current is equal to the sum of the individual branch currents. The power factor in any purely resistive circuit is one or unity and the dissipated power is equal to the square of the total current times the total resistance. The total inductance of two or more inductors connected in parallel is equal to the reciprocal of the reciprocals,

$$L_t = \frac{1}{\frac{1}{L_1}+\frac{1}{L_2}+\frac{1}{L_3}}.$$

The total capacitance of two or more capacitors connected in parallel is equal to the sum of the individual capacitances, $C_t = C_1 + C_2 + C_3$. The total reactance in either case is equal to the reciprocal of the individual reactances,

$$X_{C_t} = \frac{1}{\frac{1}{X_{C_1}}+\frac{1}{X_{C_2}}+\frac{1}{X_{C_3}}},\ X_{L_t} = \frac{1}{\frac{1}{X_{L_1}}+\frac{1}{X_{L_2}}+\frac{1}{X_{L_3}}}$$

The current leads the voltage by 90° in a purely capacitive circuit and lags the voltage by 90° in a purely inductive circuit.

The discussion of parallel circuits in this chapter is divided into discussions of resistance-inductance (RL) circuits, resistance-capacitance (RC) circuits, and resistance-capacitance-inductance (RCL) circuits. Graphic solutions and right-triangle equation solutions are applied to each type of circuit.

RL CIRCUITS

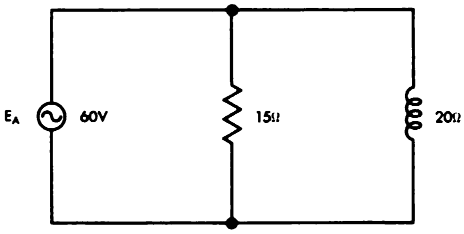
Assume that a 15-ohm resistor and an inductor whose reactance at the applied frequency is 20 ohms are connected in parallel to a 60-volt source of alternating current. Suppose that you want to find the total current.

Graphic Solution

Calculate the individual branch currents in the parallel RL circuit below as follows:

$$I_R = \frac{E_A}{R} = \frac{60}{15} = 4 \text{ amps}$$

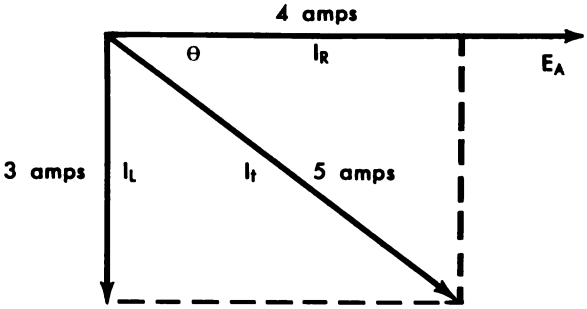
$$I_L = \frac{E_A}{X_L} = \frac{60}{20} = 3 \text{ amps}$$



Parallel RL Circuit

These currents cannot be added together because they differ in phase by 90°. The current through the resistor is in phase with the voltage across the resistor and the current through the inductor lags the voltage across the inductor by 90° (assuming a pure inductance). You can, however, find the total current by vector addition. Since the voltage is the same across all components in a parallel circuit, draw a vector representing the voltage across this circuit as the reference vector as shown in the illustration. Label this vector E_A. You need not draw this vector to scale since you are interested only in its phase or position.

Since the current through the resistor is in phase with the voltage, draw a vector representing this current along vector E_A. Label this vector I_R. The length of I_R must be determined by a convenient scale such as 1 ampere per inch. Using the same scale that you used in determining the length of I_R draw a vector representing the current through the inductor 90° clockwise from vector E_A since this current lags the voltage by 90°. Label this vector I_L. Now draw dotted lines parallel to vectors I_R and I_L as shown in the illustration. Draw in the diagonal of the parallelogram or rectangle from the center of the diagram to the opposite corner. Label this vector I_t. Measure the length of this vector and convert its length to amperes by means of the scale you used in constructing the other current vectors. Your result should be 5 amperes of total current.

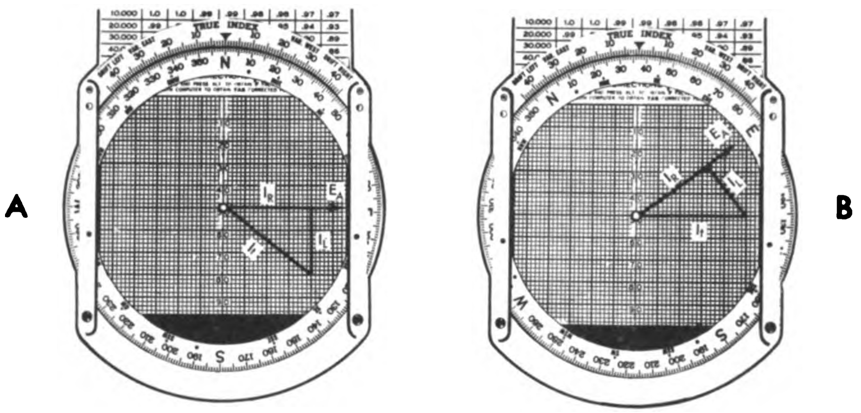


Vector Addition of Current, Parallel RL Circuit

To find the phase angle between the applied voltage and the total current, measure the angle θ between vector E_A and vector I_t with a protractor. This angle should be approximately $-37°$. In other words, the total current in this circuit lags the voltage by about 37°.

This type of problem is very easily solved on the DR computer. On the wind side put N under the true Index and the grommet over 50 on the grid. Draw in the voltage vector as a reference horizontally from the grommet. Choose a convenient scale (1 large square equals 1 amp), and draw I_R over the E_A vector, since they are in phase. Draw the I_L vector down from the I_R vector, as shown at A in the illustration on page 129, since the current through the coil lags the applied voltage by 90°.

By rotating the compass rose counterclockwise until the I_t vector is horizontal as shown at B, in the illustration, the magnitude of I_t and the phase angle by which the current lags the applied voltage may be determined. Here, $I_t = 5$ amps and $\theta = 37°$.



Determining Current and Phase Angle in a Parallel RL Circuit with the DR Computer

The total impedance of this parallel circuit can be found by substituting in the formula,

$$Z=\frac{E_A}{I_t}$$

Since the voltage is 60 volts and the total current is 5 amperes,

$$Z=\frac{60}{5}$$
$$=12 \text{ ohms}$$

Right-Triangle Solution

You can also use the right-triangle equation to solve for total current in a parallel RL circuit. However, as in series—circuit problems, this method cannot be used to find the phase angle. In solving for total current by this method, first solve for individual branch currents as before. Sketch the current triangle as shown to the right. This triangle need not be drawn to scale accurately since the result does not depend upon the dimensions of the sketch. Applying the right-triangle equation,

$$I_t=\sqrt{I_R^2+I_L^2}$$
$$=\sqrt{4^2+3^2}$$
$$=5 \text{ amperes.}$$

After finding total current by this method, you can proceed to solve for impedance as before.

Power Factor

The power factor of this circuit can be found by solving for true power and apparent power, and substituting in the formula,

$$PF=\frac{TP}{AP}$$

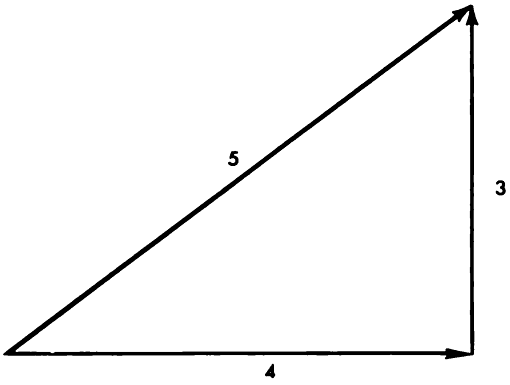
The true power consumed by the circuit is actually the power dissipated in the resistor since you have already assumed that the inductor has negligible resistance.

$$TP=I_R^2\times R$$
$$=4^2\times 15$$
$$=240 \text{ watts}$$

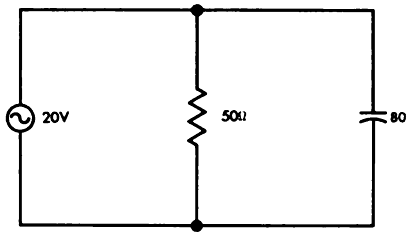
The apparent power is equal to the applied voltage times the total current

$$AP=E_A\times I_t$$
$$=60\times 5$$
$$=300 \text{ watts}$$

Substituting these values for true power and apparent power in the power factor formula you have



Current Triangle-Parallel RL Circuit



Parallel RC Circuit

$$PF = \frac{240}{300}$$
$$= .8$$

The power factor may be found more directly by using the formula,

$$PF = \cos\theta$$
$$PF = \cos 37^\circ$$
$$PF = .8$$

RC CIRCUITS

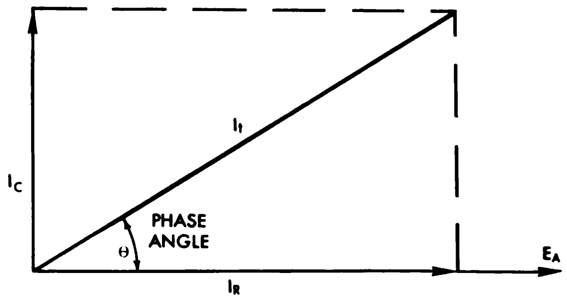
Assume that a 50-ohm resistor and a capacitor whose reactance at the applied frequency is 80 ohms are connected in parallel across a 20-volt source of alternating current, as shown above. Suppose that you want to find the total current, the phase angle, total impedance, true power, and power factor of the circuit.

Graphic Solution

To find the total current, you must first find the individual branch currents.

$$I_R = \frac{E}{R}$$
$$= \frac{20}{50}$$
$$= .4 \text{ amp}$$
$$I_C = \frac{E}{X_C}$$
$$= \frac{20}{80}$$
$$= .25 \text{ amp}$$

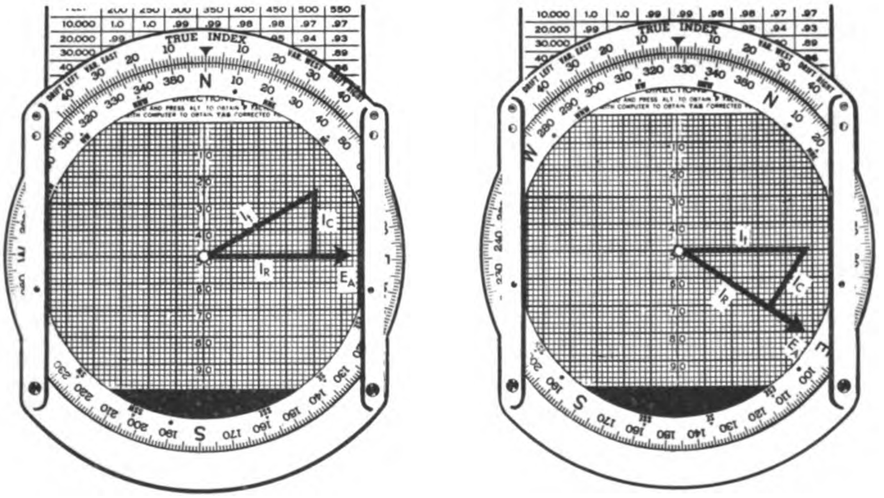
These currents must be added vectorically since the current through the resistor is in phase with the voltage and the current through the capacitor leads the voltage by 90°. Since the applied voltage is the same across both components, draw voltage vector E_A as a reference, as shown in the vector diagram. Since the current through the resistor is in phase with the voltage across the resistor, draw a vector representing this current along vector E_A. The length of this vector must be determined by a convenient scale such as .1 amp per inch. Label this vector I_R. Since the current through the capacitor leads the voltage by 90°, draw a vector to scale representing this current 90° counterclockwise from reference vector E_A. Label this vector I_C. Draw dotted lines parallel to vectors I_R and I_C as shown in the illustration. Draw in the diagonal from the center of the vector diagram to the upper right-hand corner. Label this vector I_t. Measure the length of this vector and convert its length to voltage, using the same scale that you used in constructing vectors I_R and I_C. Your result should be about .47 ampere. Measure the phase angle between the applied voltage vector E_A and total current vector I_t with your protractor. It should be approximately 32°.



Vector Additions of Currents, Parallel RC Circuit

In solving this problem on the DR computer, the procedure is the same as that used before. On the wind side, set N under the true index and the grommet over 50 on the grid. Draw E_A, the reference vector, horizontally from the grommet. Since the current through the resistor is in phase with E_A, draw the I_R vector over the reference line. In a capacitive circuit the current leads the applied voltage; therefore, I_C is drawn upward from the end of the I_R vector as shown in diagram A of the illustration on page 131.

Rotate the compass rose clockwise until the I_t vector is horizontal as shown in diagram B. The magnitude of I_t is .47 amp and the phase angle θ is 32°.



Using the DR Computer to Solve for I_t and Phase Angles in a Parallel RC Circuit

You can find the total impedance of the circuit by the formula,

$$Z_t = \frac{E_A}{I_t}$$

Since E_A equals 20 volts and the total current is .47 amp,

$$Z = \frac{20}{.47}$$
$$= 42.6 \text{ ohms}$$

Right-Triangle Method

You can also find the total current in this circuit by the right-triangle equation. First sketch a current triangle for the given conditions in the circuit and label the parts of the right triangle as shown to the right. Then apply the right-triangle formula to the parallel RC circuit as follows:

$$I_t = \sqrt{I_R^2 + I_C^2}$$
$$= \sqrt{(.4)^2 + (.25)^2}$$
$$= \sqrt{.2225}$$
$$= .47 \text{ amp}$$

Having found the total current by this method, you can find the total impedance in the same manner as before.

Power Factor

The true power dissipated in the circuit is the power consumed by the resistor, since the resistance of capacitor is so high that its effect is negligible. Since the current through the resistor is .4 ampere and the value of the resistor is 50 ohms,

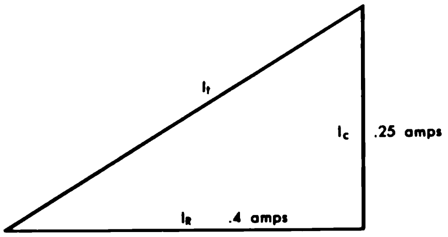
$$TP = I_R^2 R$$
$$= (.4)^2 \times 50$$
$$= 8 \text{ watts}$$

To find the power factor of the circuit, you must know the apparent power as well as the true power. Apparent power includes the power that is alternately stored by the capacitor and delivered back to the circuit. It is equal to the applied voltage multiplied by the total current. Apparent power is found as follows:

$$AP = E_A \times I_t$$
$$= 20 \times .47$$
$$= 9.4 \text{ watts}$$

Power factor can now be found by substituting the values for true power and apparent power into the power-factor formula.

$$PF = \frac{TP}{AP}$$
$$= \frac{8}{9.4}$$
$$= .85$$



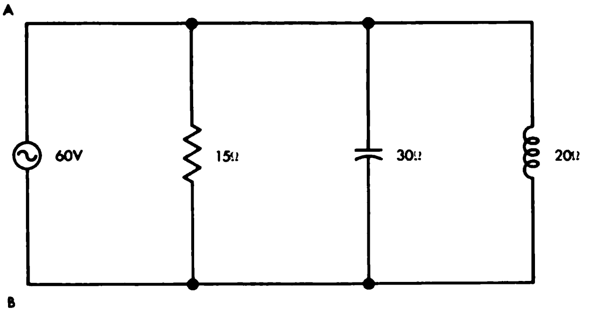
Current Triangle for Parallel RC Circuit

The power factor may be found more directly, if the phase angle is known, by using the formula

$$PF = \cos\theta$$
$$= \cos 32°$$
$$= .8480$$

RCL CIRCUITS

Assume that a 15-ohm resistor is connected in parallel with a capacitor whose reactance is 30-ohms and an inductor whose reactance is 20-ohms at the applied frequency, and that the voltage applied to this parallel circuit is 60 volts. Suppose that you want to find the total current, phase angle, total impedance, true power, and power factor of the circuit.

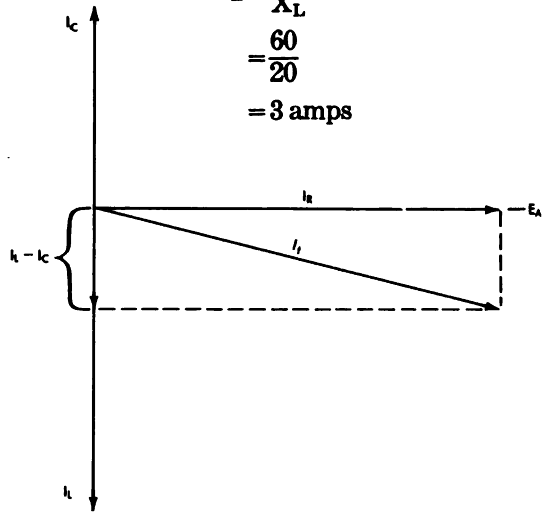


Parallel RCL Circuit

Graphic Solution

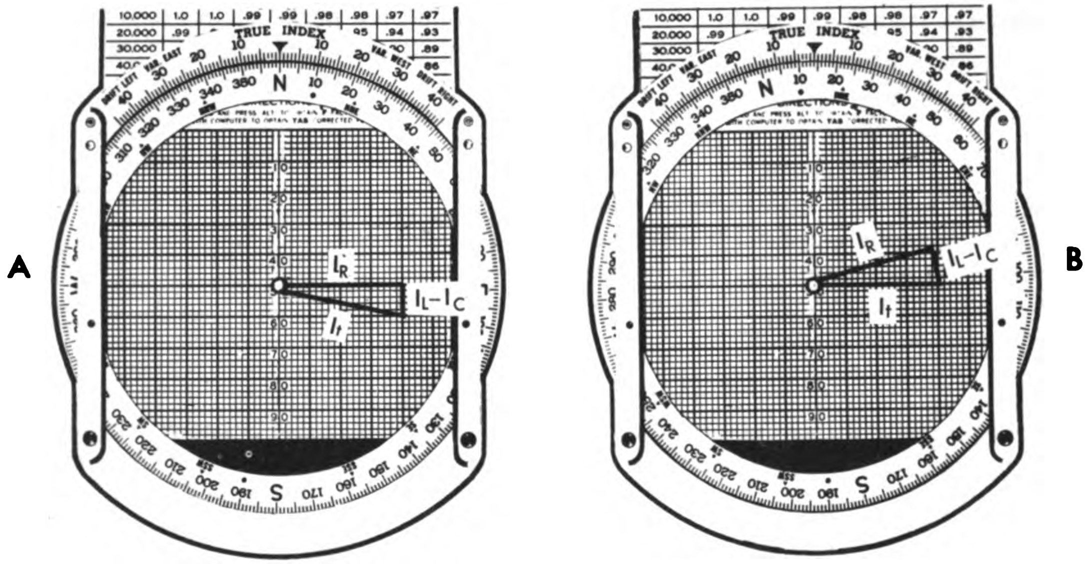
In this circuit, the current through the resistor is in phase with the voltage, the current through the capacitor leads the voltage by 90°, and the current through the inductor lags the voltage by 90°. From this information, you can see that the current flowing through the capacitor is 180° out of phase with the current flowing through the inductor. In other words at the instant that maximum current flows from A to B through the capacitor, maximum current is flowing in the opposite direction or from B to A through the inductor. The total current flowing through the circuit at any instant is, therefore, the difference of the current through the capacitor and the inductor, added vectorially to the current through the resistor. Before constructing a vector diagram of currents through the components, find the individual branch currents.

$$I_R = \frac{E}{R}$$
$$= \frac{60}{15}$$
$$= 4 \text{ amps}$$
$$I_C = \frac{E}{X_C}$$
$$= \frac{60}{30}$$
$$= 2 \text{ amps}$$
$$I_L = \frac{E}{X_L}$$
$$= \frac{60}{20}$$
$$= 3 \text{ amps}$$



Vector Addition of Current Parallel RCL Circuit

Since the voltage is the same across all three components, draw a voltage vector as a reference and label this vector E_A. Since the current through the resistor is in phase with the applied voltage, draw the current through the resistor along vector E. Label this vector I_R. The current through the capacitor leads the voltage; therefore, draw a vector representing this current 90° counterclockwise from vector E_A. Label this vector I_C. The current through the inductor lags the voltage by 90°; draw a vector representing this current 90° clockwise from vector E_A. Label this vector I_L. Be sure to draw all current vectors to scale. Subtract vector I_C from vector I_L and label this difference vector $I_L - I_C$, as shown in the vector diagram above. Form a parallelogram or rectangle by drawing dotted lines parallel to vectors I_R and $I_L - I_C$. Now draw the diagonal from the center of the vector diagram to the opposite corner. Label this vector I_t.



Using the DR Computer to Solve Parallel RCL Circuit

Measure the length of the vector and convert its length to current, using the scale that you used in constructing the other current vectors. The total current should be approximately 4.13 amperes.

Measure the angle θ between vector E_A and vector I_t with a protractor. This phase angle is approximately 14°; that is, the total current lags the applied voltage by about 14°.

To work this problem on the DR computer, proceed as follows. On the wind side turn the compass rose until N is under the true index and place the grommet over 50 on the grid. Choose a convenient scale (1 large square equals 1 ampere) and draw the I_R vector horizontally from the grommet. Find the difference of the two reactive currents ($I_L - I_C$) and plot this vector downward from the end of the I_R vector, as shown at A in the illustration shown above.

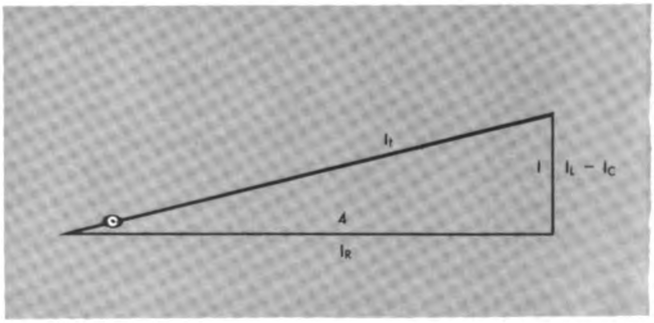
Rotate the compass rose counterclockwise until the I_t vector is horizontal and convert its length into current, 4.13 amperes. The angle by which the current lags the applied voltage is found by reading the number of degrees under the true index as shown at B. The phase angle θ is 14° for this problem.

The total impedance is equal to the applied voltage divided by the total current. Substituting 60 volts for E and 4.13 for I_t,

$$Z = \frac{E}{I_t}$$
$$= \frac{60}{4.13}$$
$$= 14.5 \text{ ohms}$$

Right-Triangle Method

You can find the total current by the right-triangle equation and then solve for the remaining unknown factors. In solving for total current by this method, sketch a current triangle from the calculated values of the individual branch currents as shown below. Subtract the current through the capacitor from the current through the inductor and make this difference value one leg of the triangle. The value of the current through the resistor becomes the other leg.



Applying the right-triangle equation,

$$
\begin{aligned}
I_t &= \sqrt{I_R{}^2+(I_L-I_C)^2} \\
&= \sqrt{(4)^2+(1)^2} \\
&= \sqrt{17} \\
&= 4.12 \text{ amperes}
\end{aligned}
$$

This value of current is a little more accurate than the value found by graphical means since the construction of the angle and the vectors introduces small errors. The impedance can be found in the same manner as before.

Power Factor

To find true power, multiply the square of the current through the resistor by its resistance.

$$
\begin{aligned}
TP &= I_R{}^2 \times R \\
&= (4)^2 \times 15 \\
&= 240 \text{ watts}
\end{aligned}
$$

To find the power factor, it is first necessary to find the apparent power.

$$
\begin{aligned}
AP &= E \times I_t \\
&= 60 \times 4.13 \\
&= 247.8 \text{ watts}
\end{aligned}
$$

Solving for power factor,

$$
\begin{aligned}
PF &= \frac{TP}{AP} \\
&= \frac{240}{247.8} \\
&= .97
\end{aligned}
$$

If the phase angle is known, the power factor may be found more directly from the formula,

$$
\begin{aligned}
PF &= \cos\theta \\
PF &= \cos 14^\circ \\
PF &= .9703
\end{aligned}
$$

Chapter 18

Resonance

You have probably observed examples of mechanical resonance many times. *Mechanical resonance* exists whenever energy is being applied to a vibrating object at a frequency equal to its natural frequency of vibration. For example, consider the pendulum of a pendulum-type clock. A large clock of this type has a seconds pendulum that is slightly more than 39 inches long. One second is required for a pendulum of this length to swing from one side to the other. The resistance of the air and the pull of gravity tend to reduce the back and forth swing of the pendulum; therefore, some energy must be applied to it during each swing so that the clock will continue to run. This is accomplished by a ratchet-like mechanism called the escapement. The escapement releases energy stored in the mainspring at the correct time during each swing to give the pendulum a push in the direction that it is starting to swing. You can say, therefore, that the pendulum is mechanically resonant to the pulses of energy that are released from the mainspring. You can also say that the *resonant frequency* of the pendulum, in this case, is one-half cycle per second, since it requires two seconds to swing from one side to the other and back.

An electric circuit can possess the property of *electrical resonance*. In order for a circuit to be resonant, it must have both inductance and capacitance. A capacitor stores energy during one half of the cycle of the applied alternating voltage and delivers this energy back to the circuit during the other half of the cycle. This energy is stored in the electric field that exists between the plates of the capacitor. An inductor also stores energy during one half cycle and returns the energy during the next half cycle. Energy is stored in the magnetic field of an inductor. However, the capacitor and inductor do not store energy during the same half cycle. When the capacitor is storing energy, the inductor is giving up energy and vice versa. This makes it possible for the inductor and the capacitor to pass the same energy back and forth. The rate at which this energy is passed back and forth can be called the *resonant frequency* of the circuit. More specific definitions of resonance are given for series and parallel resonant circuits later in the chapter.

OPERATION OF A RESONANT CIRCUIT

One complete chain of events that takes place in a resonant circuit is illustrated on page 137. Assume that a charge is placed across the capacitor as shown in 1. In 2, the switch is closed to the right to complete the circuit between the capacitor and the inductor and places the inductor in series with the capacitor. (The following discussion applies to the operation of both series-tuned and parallel-tuned circuits.) This furnishes a path for the excess electrons on the upper plate to flow to the lower plate and thus neutralize the charge. As current flows through the coil, a magnetic field is built up around the coil. In this manner the electric energy that was first stored by the capacitor is now stored in the magnetic field of the inductor. In 3, the charge of the two plates is shown equalized and a maximum stationary magnetic field exists around the coil. Since the capacitor is completely discharged, the magnetic field starts to collapse (4). The collapsing field induces a voltage in the coil which tends to cause the current to continue to flow toward the bottom plate. At 5, the magnetic field has completely collapsed, but in doing so it has caused the capacitor to become charged with the opposite polarity to that of the original charge in 1. If the circuit has no resistance, the amount of this charge would be the same as the original charge. However, the coil and the connecting wires have some resistance, and a small amount of energy is dissipated in the form of heat (I^2R loss). Therefore, the charge in 5 is slightly less than the original charge. The capacitor can now start to discharge back up through the coil. This discharge current (6) causes a new magnetic field to be built up around the coil. At 7, the charge on the two capacitor plates has equalized and all of the energy is stored in the magnetic field again. The magnetic field starts to collapse in 8, causing further electron flow toward the upper capacitor plate. This action is completed in 9, and the capacitor now has the same polarity of charge that it had in 1. During this second exchange of energy, even more energy is dissipated in the form of heat so that the capacitor charge in 9 is even less than the capacitor charge in 5.

If the circuit is not opened, this discharging and charging action will continue until all of the energy in the circuit has been dissipated as heat. The number of times that this complete set of events occurs per second is called the natural frequency or the *resonant frequency* of the circuit.

If it were possible to throw the switch over to battery quickly after each complete set of events (as at time 9) to replenish the dissipated energy and to complete the DC circuit again, the action would continue indefinitely. This switching action can be accomplished by means of an electron tube. When this is done, you can say that the circuit is resonant to the same frequency as the switching action.

SERIES RESONANCE

A series resonant circuit is a series *circuit in which the inductive reactance equals the capacitive reactance.* This condition may be expressed in formula form as:

$$X_L = X_C$$

Resonant Frequency

A formula for resonant frequency may be derived from this definition of a series resonant circuit. Since $X_L = 2\pi f L$ and $X_C = \frac{1}{2\pi f C}$

$$2\pi f L = \frac{1}{2\pi f C}$$

Solving this equation for f, you can derive a formula for finding the resonant frequency. Multiplying both sides of the equation by $2\pi f C$,

$$4\pi^2 f^2 LC = 1$$

Dividing both sides of the equation by $4\pi^2 LC$,

$$f^2 = \frac{1}{4\pi^2 LC}$$

Taking the square root of both sides of the equation,

$$f = \frac{1}{2\pi\sqrt{LC}}$$

where f is in cycles per second, L is in henries, and C is in farads. This formula can be used to find either f, L, or C, if the other two values are known.

Example 1

A .265-henry inductor is connected in series with a 26.5-mfd capacitor. Find the resonant frequency.

Solution:

Substitute .265 for L and 26.5×10^{-6} for C in the formula for series resonance and solve.

$$f = \frac{1}{6.28\sqrt{.265 \times 26.5 \times 10^{-6}}}$$

$$= \frac{.159}{\sqrt{7.0225 \times 10^{-6}}}$$

$$= \frac{.159}{2.65 \times 10^{-3}}$$

$$= 60 \text{ cycles per second}$$

Example 2

How much capacitance is necessary to resonate a .25-millihenry coil at a frequency of 1000 kc?

Solution

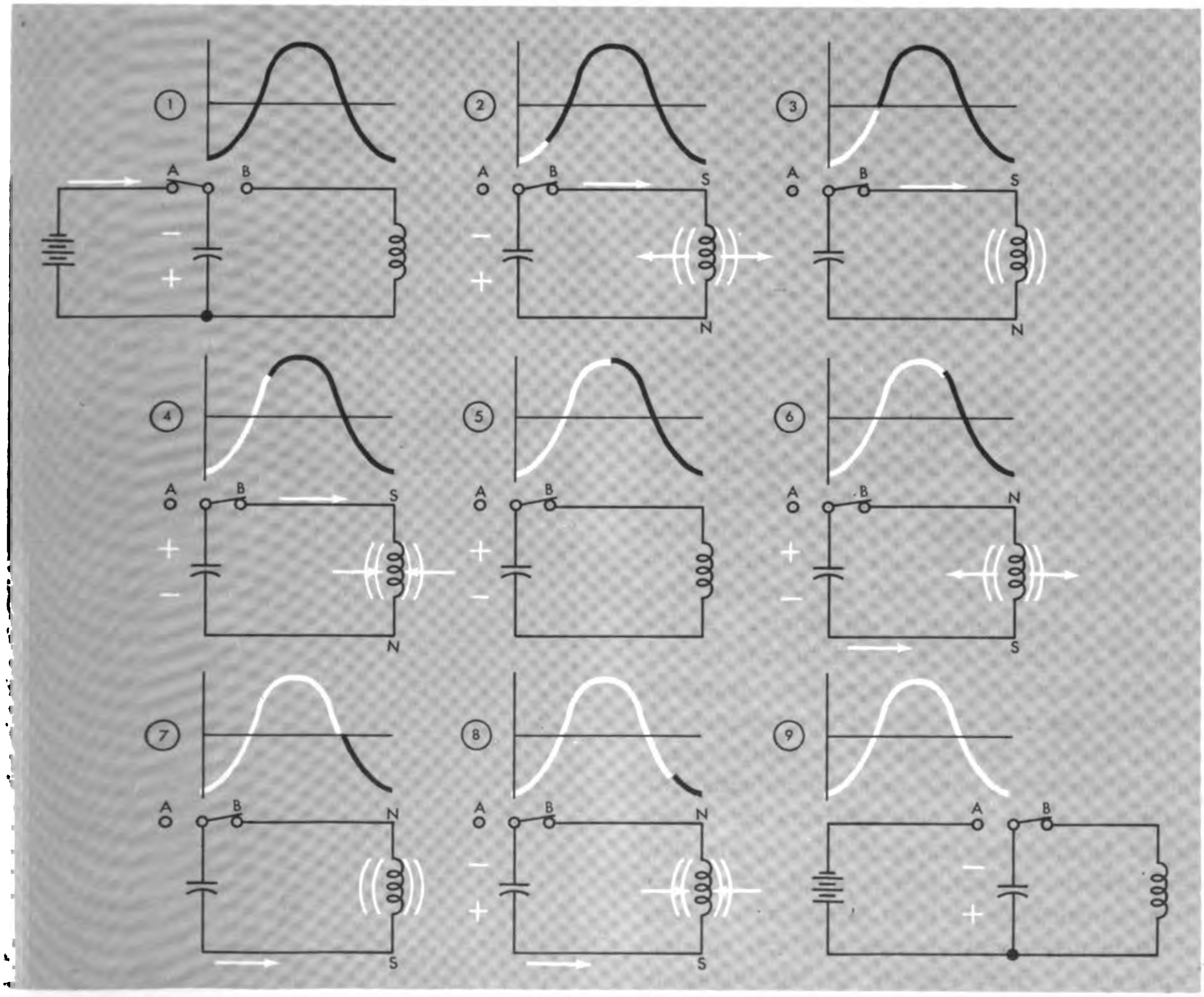
Solve the formula for resonant frequency for C.

$$f = \frac{1}{2\pi\sqrt{LC}}$$

Squaring both sides of the equation,

$$f^2 = \frac{1}{4\pi^2 LC}$$

Dividing both sides by f^2 and multiplying by C,



Electrical Resonance

$$C=\frac{1}{4\pi^2f^2L}$$

Substituting 1000 kc or 10^6 for f, and $.25\times10^{-3}$ for L,

$$C=\frac{1}{39.44\times10^{12}\times.25\times10^{-3}}$$

$$=\frac{1}{9.86\times10^{9}}$$

$$=101.4 \text{ mmf}$$

Current Is Maximum

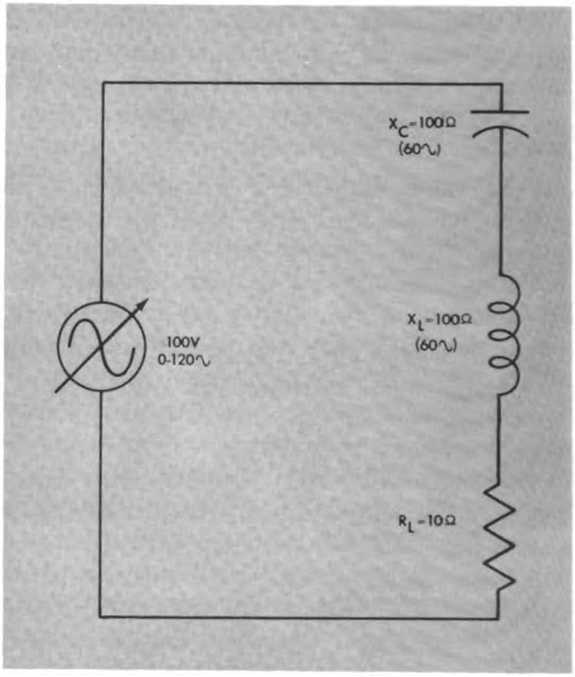
Study the circuit in the illustration to the right. The frequency of the alternator can be varied from 0 to 120 cps with a constant output voltage of 100 volts. The resistance of the inductor and connecting wire, as measured by an ohmmeter, is 10 ohms. When the output frequency of the alternator is 60 cps, the inductive reactance and the capacitive reactance are each equal to 100 ohms. This frequency is, therefore, the resonant frequency of the circuit. Since these two reactances oppose each other, the total opposition to the flow of current is 10 ohms. Solving for current flow,

$$I=\frac{E}{R}$$

$$=\frac{100}{10}$$

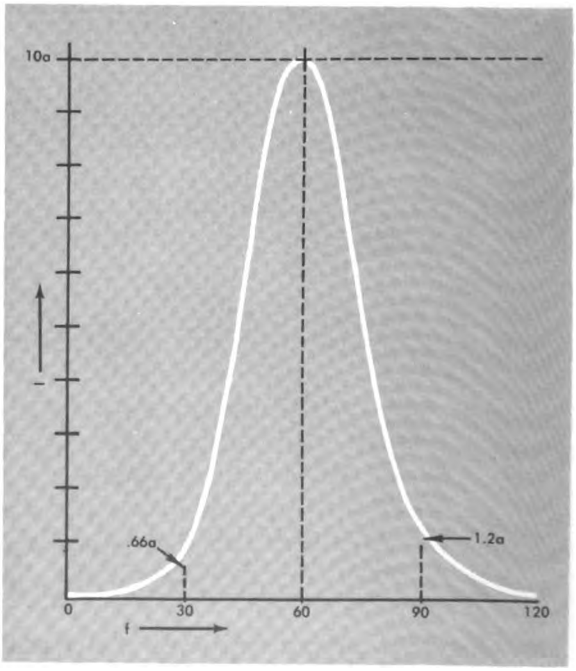
$$=10 \text{ amperes}$$

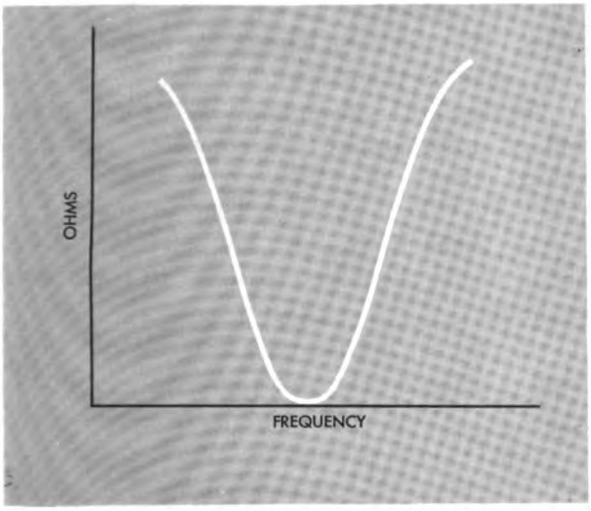
If the applied frequency is decreased, the capacitive reactance becomes greater and the inductive reactance becomes less. For example, at 30 cps the capacitive reactance is 200 ohms and the inductive reactance is 50 ohms. Solving for total current at 30 cps, you will find that I_t is approximately .66 ampere. If the applied frequency is increased above the resonant frequency, the capacitive reactance becomes less and the inductive reactance becomes greater. At 90 cps the capacitive reactance is about 70 ohms and the inductive reactance is about 150 ohms. Solving for current at an applied frequency of 90 cps, you will find that the current flow is about 1.2 ampere. You can calculate the total current at a number of frequencies between zero and 120 cps and construct a curve similar to the resonance curve for the series circuit shown to the right. Note that *current is maximum* at resonance.



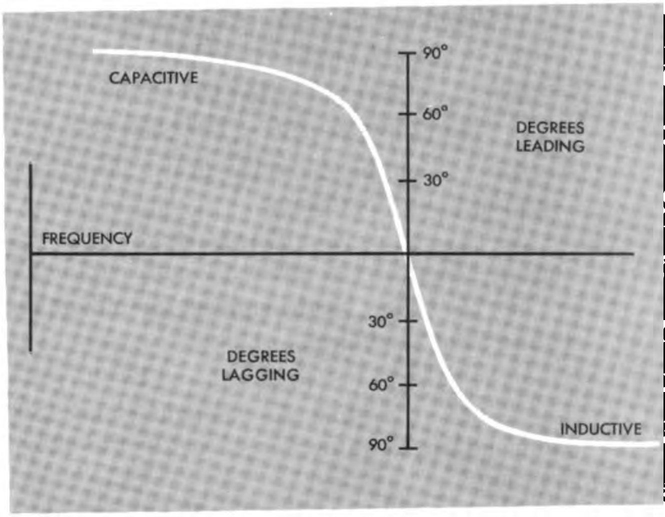
Series LCR Circuit

Resonance Curve for Series LCR Circuit





Impedance Curve for Series LCR Circuit



Phase Angle of Current for Series LCR Circuit

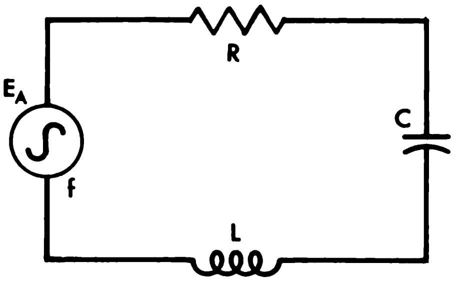
Impedance Is Minimum and Resistive

Since current is maximum at resonance, impedance must be minimum. This fact may be deduced from Ohm's law for AC: $Z = \frac{E}{I}$. Impedance decreases as current increases; therefore, impedance is minimum when current is maximum. The impedance curve shown in the illustration above is an inversion of the resonance curve.

A series LC circuit acts like a capacitance below the resonant frequency since the capacitive reactance is greater than the inductive reactance and, therefore, more voltage is developed across the capacitor. Consequently, the current leads the applied voltage. If the voltage applied to the series LC circuit has a frequency higher than the circuit's resonant frequency, the circuit acts like an inductance. In this case the current lags the applied voltage. The change from leading to lagging current is shown in the illustration to the right above. At the resonant frequency, the circuit acts like a very low value of resistance. This fact is evident from the formula for impedance which you have already studied,

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

Since $X_L = X_C$ at resonance, the value of $X_L - X_C$ is equal to zero, and $Z = R$. Therefore, the impedance of a series resonant circuit is *minimum and resistive.*



Power Factor Is Unity

Another indication of series resonance is that the *power factor is equal to one.* You have learned that the formula for power factor may be expressed as:

$$PF = \frac{R}{Z}$$

Since $R = Z$ at resonance, $PF = 1$.

VECTOR DIAGRAM ANALYSIS

The conclusions reached in the preceding discussion can also be reached through the use of vector diagrams. We will begin by constructing the vector diagrams for a series circuit AT resonance.

The figure below shows a series RLC circuit in which the applied frequency is equal to the resonant frequency. The *impedance* and *voltage-current* vector diagrams can easily be drawn from information with which you are already familiar.

The *impedance* diagram is as shown below.

1. R is the reference vector.
2. X_L is always plotted 90° ahead of R.
3. X_C is always plotted 90° behind R.
4. Z_T can be correctly drawn using the following line of reasoning.

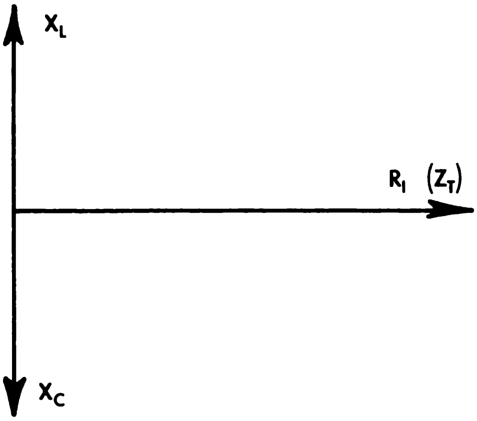
a. X_L equals X_C. (This is one of the conditions or criteria for a series resonant circuit.)

b. Z_T is the vector sum of the circuit resistance and reactances.

$$\vec{Z_T} = \vec{R} + \vec{X_L} + \vec{X_C}$$

c. Since X_L and X_C are equal and 180° out of phase, these factors cancel in the above equation.

d. Z_T is therefore equal to R AT resonance.



Z Diagram (AT Resonance)

The next diagram shows the *voltage-current* vector diagram for the basic resonant circuit.

1. I_T is the reference vector for all *series* circuits.
2. E_L *leads* I_T by 90° (the E across a coil leads the I through it by 90° and I_T flows through each component in a simple series circuit).
3. E_C *lags* I_T by 90°.
4. E_R is in phase with I_T (the E across a resistor is always *in phase* with the I through it; again, I_T flows through each component in a simple series circuit).
5. E_A can be correctly drawn using the following line of reasoning (we are interested in the correct placement of E_A because the angular relationship of E_A to I_T — leading or lagging and the size of the angle — is indicative of how the circuit is acting).

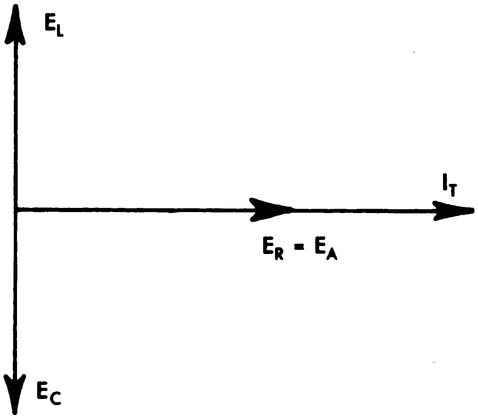
a. E_L equals E_C (this is one of the conditions for series resonance).

b. E_A is the vector sum of the individual voltage drops around the circuit (Kirchhoff's voltage law).

$$\vec{E_A} = \vec{E_R} + \vec{E_L} + \vec{E_C}$$

c. Since E_L and E_C are equal and 180° out of phase, these factors cancel in the above equation.

d. E_A is therefore equal to E_R AT resonance.



E-I Diagram (AT Resonance)

Now let's consider the vector diagrams when the frequency applied to the basic RLC circuit is *above* the resonant frequency. Begin with the *impedance* diagram on the next page.

1. R is again the reference vector.
2. X_L is greater than X_C. (If X_L equals X_C AT resonance and the new frequency is greater than the resonant frequency, it follows that X_L is now greater than X_C.)

$$\uparrow X_L = 2\pi f \uparrow L$$

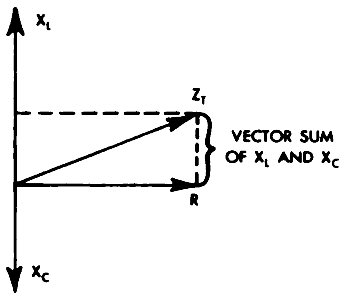
$$\downarrow X_C = \frac{1}{2\pi f \uparrow C}$$

3. The placement of Z can be determined as follows:

a. $\vec{Z_T} = \vec{R} + \vec{X_L} + \vec{X_C}$

b. Adding X_L and X_C vectorially, yields a vector which lies along the X_L axis but is shorter than X_L.

c. Z_T then equals the vector sum of R and the new vector. This shows that Z_T *leads* R and that the circuit is inductive.



Z Diagram (ABOVE Resonance)

We can arrive at the same conclusion from the *voltage-current* diagram below.

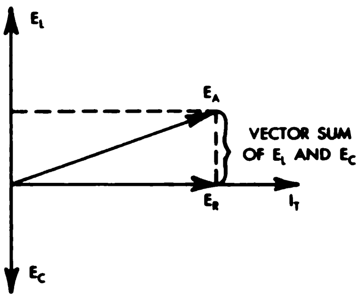
1. I_T is the reference vector.
2. E_L is greater than E_C.
3. The placement of E_A can be determined as follows:

a. E_A is the vector sum of the individual voltage drops around the circuit (Kirchhoff's voltage law).

$$\overrightarrow{E_A}=\overrightarrow{E_R}+\overrightarrow{E_L}+\overrightarrow{E_C}$$

b. First, adding E_L and E_C yields a vector along the E_L axis but shorter than E_L.

c. E_A is then the vector sum of E_R and the new vector. This shows that I_T *lags* E_A; therefore, the circuit is inductive.



E-I Diagram (ABOVE Resonance)

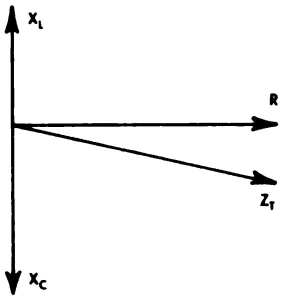
Finally, we will discuss the vector diagrams when the frequency applied to the basic RLC circuit is *below* the resonant frequency. First, consider the *impedance* diagram at top right.

1. R is the reference vector.
2. X_C is now greater than E_L:

$$\downarrow X_L=2\pi f\downarrow L$$

$$\uparrow X_C=\frac{1}{2\pi f\downarrow C}$$

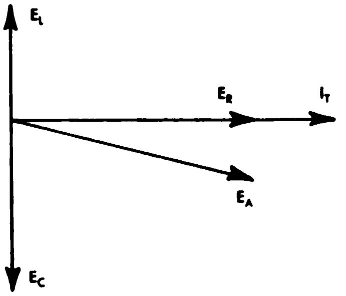
3. $\overrightarrow{Z_T}=\overrightarrow{R}+\overrightarrow{X_L}+\overrightarrow{X_C}$
4. Adding X_L and X_C yields a vector along the X_C axis but shorter than X_C.
5. Adding the new vector to R yields Z_T. The diagram shows that Z_T *lags* R; therefore, the circuit is capacitive.



Z Diagram (BELOW Resonance)

The *voltage-current* diagram for the circuit *below* resonance is illustrated below.

1. I_T is the reference vector.
2. E_C is greater than E_L.
3. $\overrightarrow{E_A}=\overrightarrow{E_R}+\overrightarrow{E_L}+\overrightarrow{E_C}$ (Kirchhoff's voltage law)
4. The vector sum of E_L and E_C yields a vector along the E_C axis but shorter than E_C.



E-I Diagram (BELOW Resonance)

5. E_A is then the vector of sum of E_R and the new vector. The diagram shows that with the circuit *below* resonance, I_T *leads* E_A; therefore, the circuit is capacitive.

Vector analysis shows, then, that *at* resonance, every *series* RLC circuit is *resistive; above* resonance, every *series* RLC circuit is *inductive;* and *below* resonance, every *series* RLC circuit is *capacitive.* Although the *above, below,* and *at* resonance conditions were attained by changing frequency, a change in L or C would have given the same result.

Q of a Tuned Circuit

In multiplying the current by the inductive reactance or the capacitive reactance at resonance, you can find that the voltage across either reactive component in the circuit on page 138 is about 1000 volts. Since the applied voltage is only 100 volts, the voltage across one of the reactive components represents a *voltage gain* of 10. This property of series-tuned circuits makes them very useful in electronics and communications equipment.

The amount of gain possible in a tuned circuit depends upon the relative amounts of inductive reactance and resistance in the circuit. This ratio is called the figure of merit or Q, and is expressed as follows:

$$Q = \frac{X_L}{R}$$

In the circuit on page 138 the inductive reactance is 100 ohms at a frequency of 60 cps and the resistance of the coil is 10 ohms. Substituting these values in the above formula,

$$\begin{aligned} Q &= \frac{X_L}{R} \\ &= \frac{100}{10} \\ &= 10 \end{aligned}$$

If the resistance of the circuit with an applied voltage of 100 volts were 20 ohms, the Q would have been 5. The voltage across the coil at resonance would then have been 500 volts. On the other hand, if the resistance of the coil were 5 ohms, the Q would have been 20. The voltage across the coil at resonance would have been 2,000 volts with an applied voltage of 100 volts.

It is usually desirable to have a resonant circuit with as high a value of Q as possible in order to realize the maximum amount of gain from the circuit. A Q of 20 or more is considered high, and a Q of 200 to 400 is obtainable with ferrite cores.

The Q of the circuit is practically the same as the Q of the coil. The resistance of the connecting wires is small in comparison to the DC resistance of the coil itself.

Bandwidth

Consider the series LC circuit whose resonance curve is shown in the illustration to the right. As the frequency is varied through resonance, the current increases from some low value to a maximum and then decreases to a low value again. The current flow is maximum at the resonant frequency and diminishes *slightly* at frequencies just below and just above resonance. However, the current flow is greatly diminished at frequencies far below and far above resonance. Those frequencies at which current flow is comparatively great are called a *band* of frequencies. The *bandwidth* is the difference between the highest and lowest frequencies contained in the band. Those frequencies just outside the band allow almost as much current to flow as those just inside the band. What determines the width of the band? Where shall the dividing line be? A standard that has been universally adopted states that the limiting frequencies which mark the two edges of the band are those at either side of resonance at which the current falls to *0.707 of the current at resonance*. These are known as the half-power points (a and b of the upper illustration on page 143).

The fact that the current at the half-power points is equal to 0.707 of the current at resonance may be shown as follows:

At resonance $P = I^2R$

At half-power point $P_1 = I_1{}^2R$

Since $P_1 = \frac{P}{2}$

$$I_1{}^2R = \frac{I^2R}{2}$$

Dividing by R_1

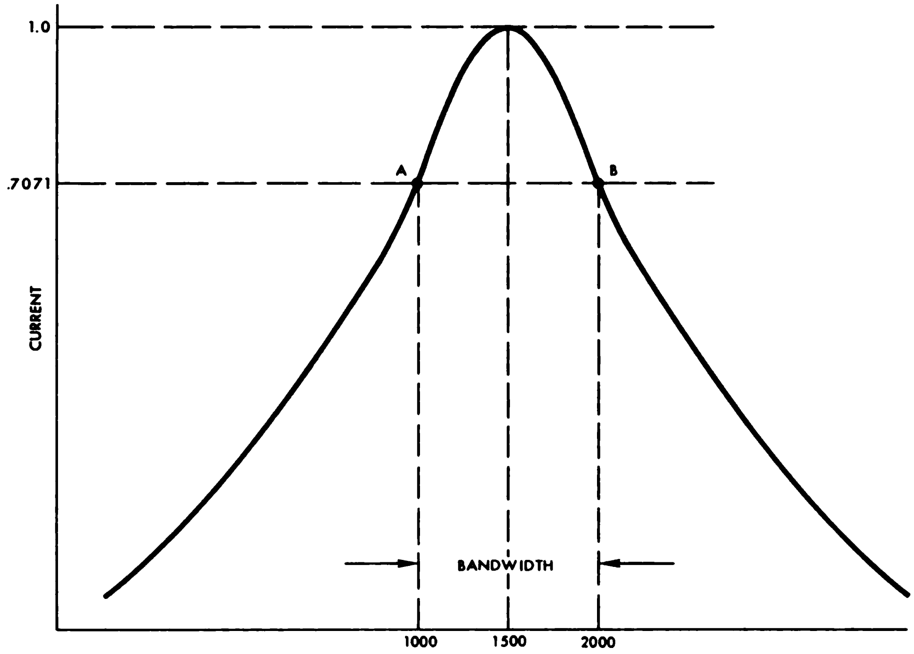
$$I_1{}^2 = \frac{I^2}{2}$$

Taking the square root of both sides

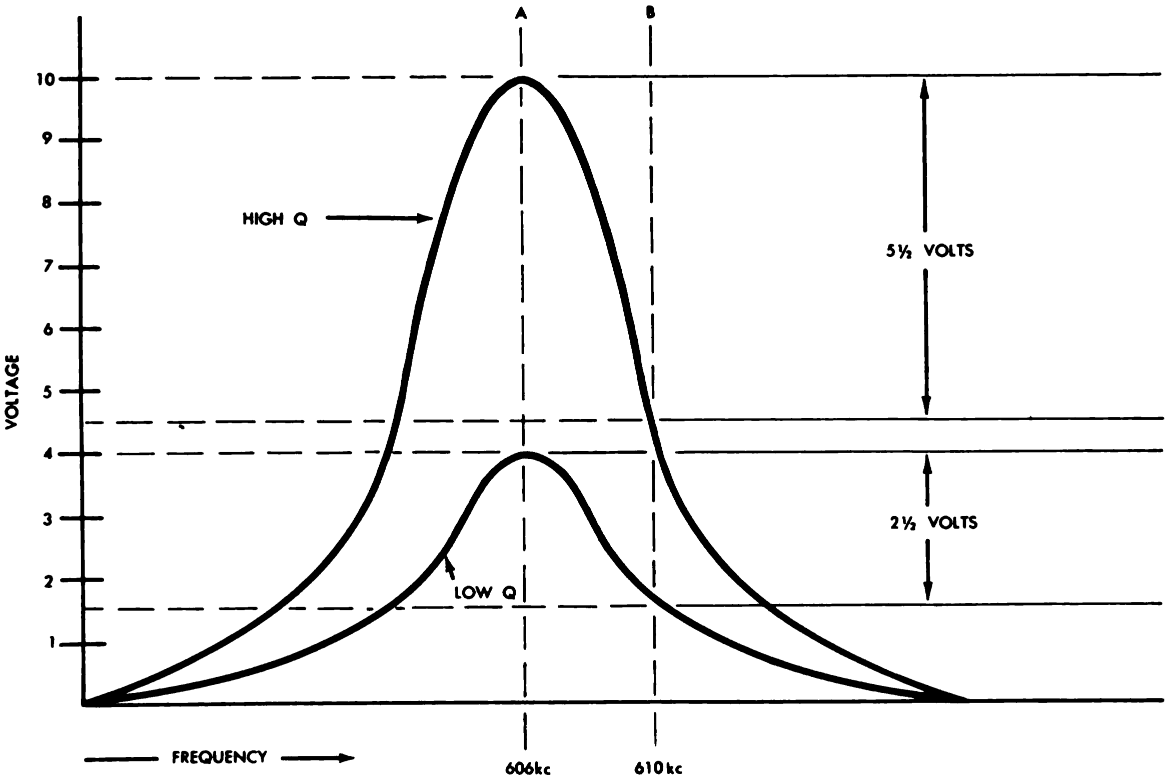
$$\begin{aligned} I_1 &= \frac{I}{\sqrt{2}} \\ &= \frac{I}{1.414} \\ &= 0.7071 \end{aligned}$$

Selectivity

The Q of a resonant circuit determines the circuit's *selectivity*. The selectivity of a resonant circuit is the ability of the circuit to discriminate



Bandwidth of a Series LCR Circuit



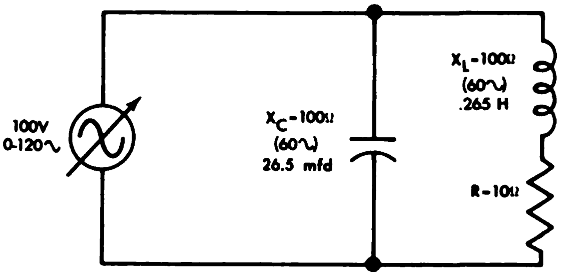
Effect of Q on Selectivity

between several frequencies that are close together. For example, assume that two voltages at two different frequencies are induced into the coil of an LC circuit from another coil. Call these frequencies A and B. Suppose that frequency A is the resonant frequency of the circuit. Both induced voltages tend to cause circulating currents to flow. That is, both cause energy to be stored alternately in the capacitor and in the inductor. As a result of these circulating currents, a usable amount of voltage is developed across the coil or the capacitor at the resonant frequency A, but practically no voltage is developed at frequency B. You can then say that the circuit has discriminated against frequency B.

The steepness of the resonance curve indicates the amount of selectivity of the circuit. If the circuit has a high Q and, therefore, a steep resonance curve, its selectivity is good. If there is a relatively large amount of resistance in the circuit, the resonance curve will not be very steep and the selectivity will be poor. This is shown in the lower illustration on page 143. The difference in voltages developed at frequencies A and B in the high-Q circuit is 5½ volts. The difference in voltages developed at the same two frequencies by the low-Q circuit is only 2½ volts.

PARALLEL RESONANCE

The most common condition of parallel resonance occurs when the inductive reactance and the capacitive reactance are equal; that is, $X_L = X_C$. This is the same condition that was given for series resonance; therefore, the same formula for resonant frequency can be used for series and parallel resonance



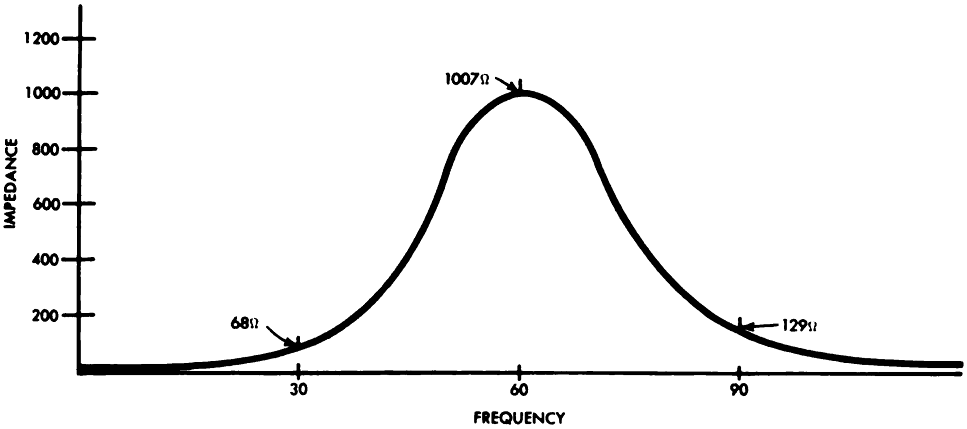
Parallel LC Circuit

$$f = \frac{1}{2\pi\sqrt{LC}}$$

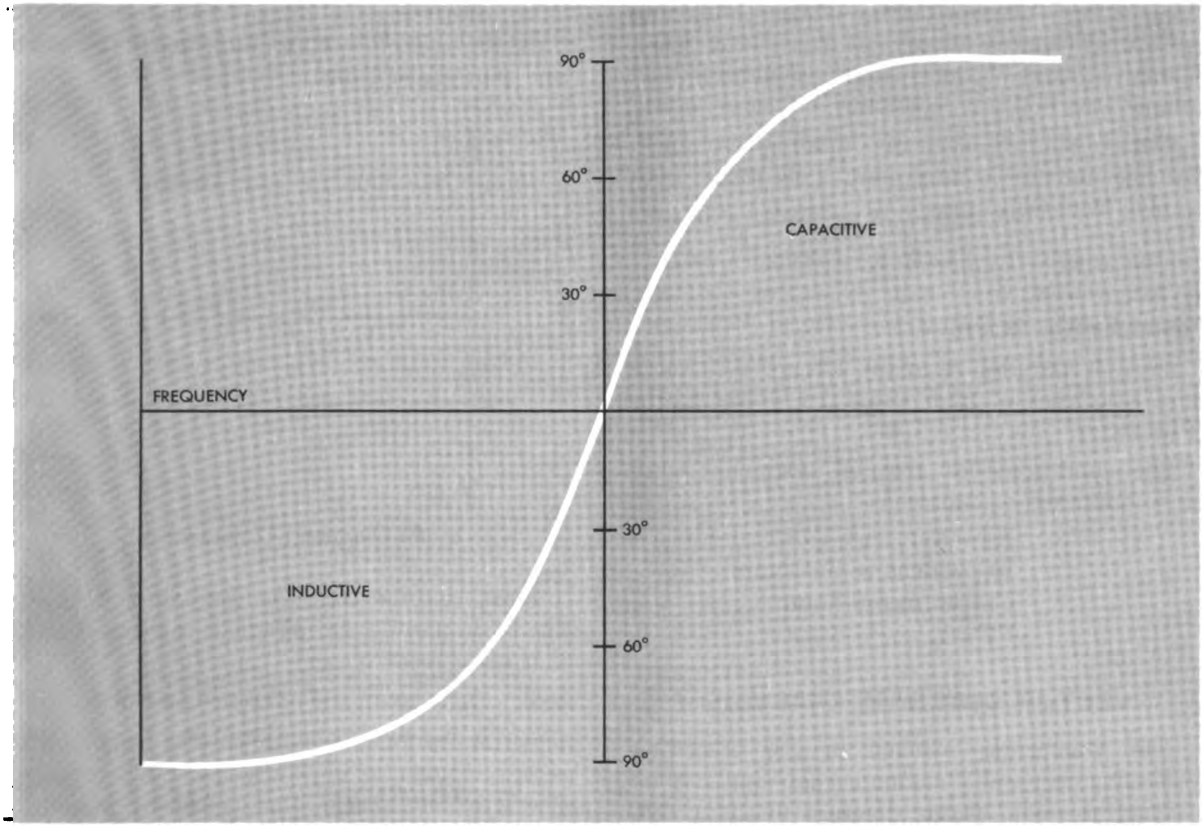
Impedance Is Maximum and Resistive

In the circuit shown, a .265-henry inductor is connected in parallel with a 26.5-mfd capacitor to a source of alternating current. The frequency of this source is variable from zero to 120 cps, but the output voltage is constant at 100 volts. The total resistance of the inductive branch of the circuit is 10 ohms and is represented by resistor R. Most of this resistance is the resistance of the coil. The reactances of both inductor and capacitor equal 100 ohms at a frequency of 60 cps. This is, therefore, the resonant frequency of the circuit.

The characteristics of a parallel-resonant circuit can be more readily understood with the aid of an *impedance curve.* The impedance curve shown below was constructed from the conditions of the parallel circuit illustrated above. In constructing a curve of this type, you calculate



Impedance Curve for Parallel LC Circuit



Phase Angle of Current for Parallel Circuit

the impedance of the circuit for a number of frequencies and plot these impedance values on a graph. When you have connected all of the points with a smooth curve, you have an impedance curve. With respect to this particular circuit, the impedance at the resonant frequency of 60 cps is about 1,007 ohms. At 30 cps the impedance is about 68 ohms and at 90 cps the impedance is about 129 ohms.

At frequencies below 60 cps the capacitive reactance is greater than the inductive reactance. Since more current is permitted to flow through the inductor, you can say that the circuit acts like a series inductor, and the total current in the circuit lags the applied voltage by some angle between zero and 90°. Above resonance the capacitive reactance is less than the inductive reactance, and the circuit acts like a series capacitance. Consequently, the total current or line current leads the applied voltage. The phase angle of current is shown in the illustration above. At the resonant frequency, the circuit is neither inductive nor capaci tive, but acts like a large value of resistance At resonance, therefore, the line current and the applied voltage are in phase.

Current Is Minimum

Since the current leads the voltage by 90° in the capacitive branch and lags the voltage by 90° in the inductive branch at resonance, the two currents flow in opposite directions. That is, when the capacitor is discharging, the discharge current flows through the inductor, storing energy in its magnetic field. At the time when the magnetic field is collapsing, the resulting current flows into the capacitor, storing energy in the electric field. If the parallel circuit had no losses, the voltage across it would always equal the applied voltage and no current could flow from the alternator (line current). The impedance of the circuit would then appear to be infinite. However, no circuit exists that does not have at least a small amount of resistance and, therefore, some of the circulating current dissipates energy as heat (I^2R loss). A small amount of line current is then permitted to flow.

Since the impedance is maximum at resonance, the line current is minimum. This result may be seen from Ohm's law

$$I = \frac{E}{Z}$$

Power Factor Is Unity

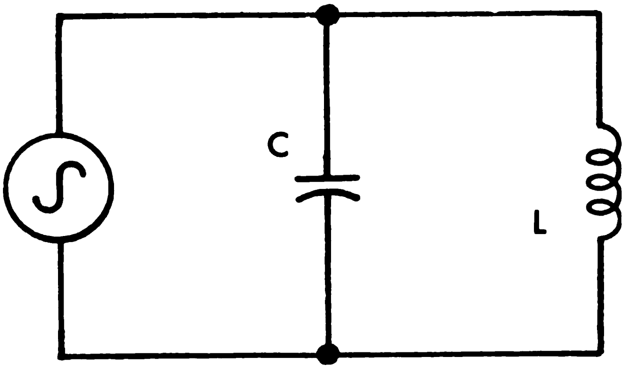
Since a parallel resonant circuit acts as a pure resistance of high value, its power factor is unity.

The conditions for parallel resonance may be summarized as follows:

$X_L = X_C$
Impedance is maximum
Current is minimum
$PF = 1$

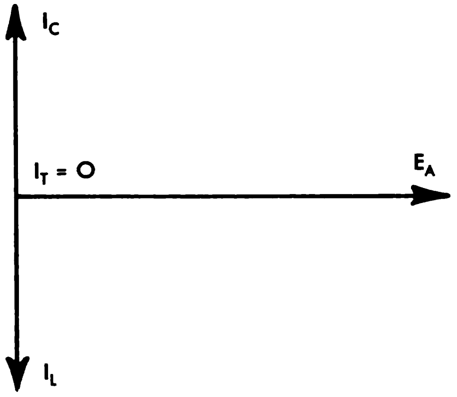
VECTOR DIAGRAM ANALYSIS

The conclusions reached in the preceding discussion on parallel resonance can similarly be reached through vector diagram analysis. We will begin with the *voltage-current* diagram for a simple parallel circuit in which the frequency applied is equal to the resonant frequency.



Parallel LC Circuit

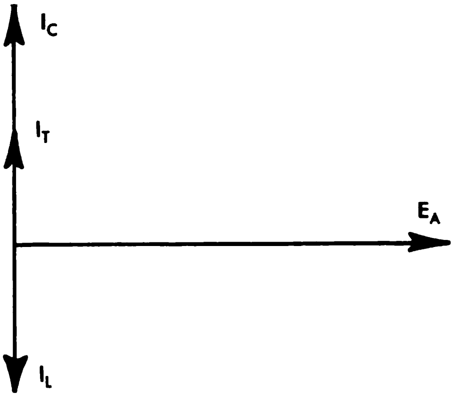
1. E_A is the reference vector.
2. I_C is drawn 90° *ahead* of E_A (I_C *leads* E_C by 90° and E_C is equal to E_A).
3. I_L is drawn 90° *behind* E_A.
4. I_T is the vector sum of I_L and I_C (Kirchhoff's current law).
5. Since I_L equals I_C, I_T equals zero (true only for a circuit containing pure inductance and pure capacitance).



E-I Diagram (AT Resonance)

Next, assume that the frequency applied to the basic parallel LC circuit is increased so that it is *above* resonance.

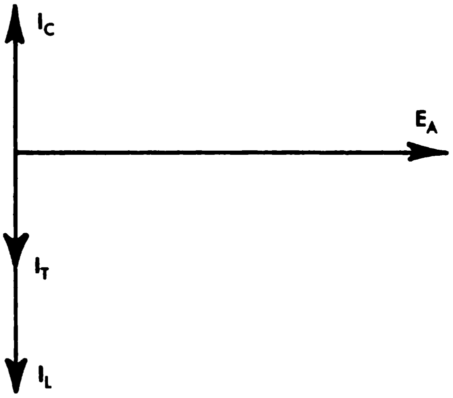
1. E_A is again the reference vector.
2. I_C is now greater than I_L.
   a. $\downarrow X_C = \frac{1}{2\pi f \uparrow C}$
   b. $\uparrow X_L = 2\pi f \uparrow L$
   c. $\uparrow I_C = \frac{E_A \rightarrow}{X_C \downarrow}$
   d. $\downarrow I_L = \frac{E_A \rightarrow}{X_L \uparrow}$
   e. I_C has therefore increased while I_L has decreased.
3. I_T is the vector sum of I_L and I_C. Adding these vectors yields I_T along the I_C axis. Since I_T *leads* E_A, the circuit is *capacitive*.



E-I Diagram (ABOVE Resonance)

Finally, consider the condition when the frequency applied is decreased to some frequency *below* resonance.

1. E_A is the reference vector.
2. I_L is greater than I_C.
3. I_T therefore *lags* E_A and the circuit is *inductive*.



E-I Diagram (BELOW Resonance)

Thus the *parallel* circuit *above* resonance is opposite to the *series* circuit in that the former is always *capacitive* and the latter is always *inductive*. Likewise, the *parallel* circuit *below* resonance is always *inductive* while the *series* circuit is always *capacitive*.

Effects of Q

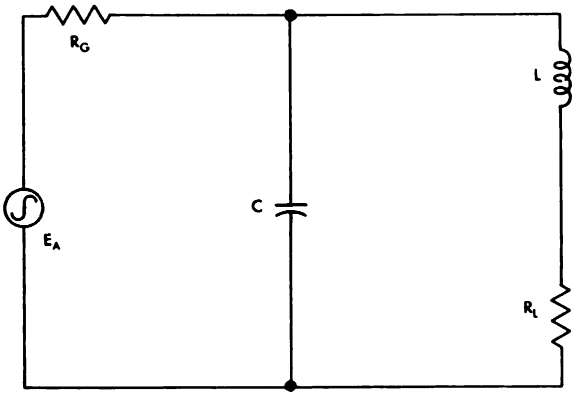
If the Q of a parallel LC circuit is low (less than 10), each of the conditions for parallel resonance exists at a *slightly* different frequency. If the parallel LC circuit has a high Q (greater than 10), these conditions exist at practically the same frequency. Since resonant or tuned circuits used in communications equipment are of the high-Q type, all conditions of resonance may be considered to occur at the same frequency, $f = \frac{1}{2\pi\sqrt{LC}}$

Bandwidth and Selectivity

The concepts of bandwidth and selectivity for parallel resonant circuits are similar to those for series resonant circuits — namely, that the narrower the bandwidth of the resonance curve, the better the circuit selectivity. However, instead of using the value of .707I for determining the half-power points, impedance (or voltage) values are used. The impedance (or voltage) at the half-power point is $Z_1 = .707Z$ (or $E_1 = .707E$) where Z_1 and E_1 are the impedance and voltage at the half-power point and Z and E are the corresponding values at resonance.

Factors Affecting Selectivity

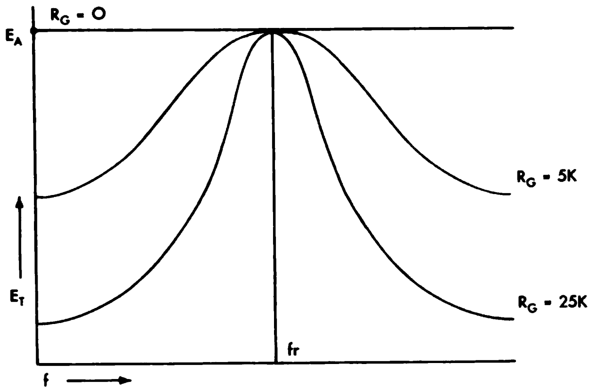
The selectivity of parallel resonant circuits depends primarily upon two factors — the resistance in series with the parallel tank and the internal resistance of the coil. (These two factors are labeled R_G and R_L in the figure below, where R_G is the internal resistance of an AC voltage source.)

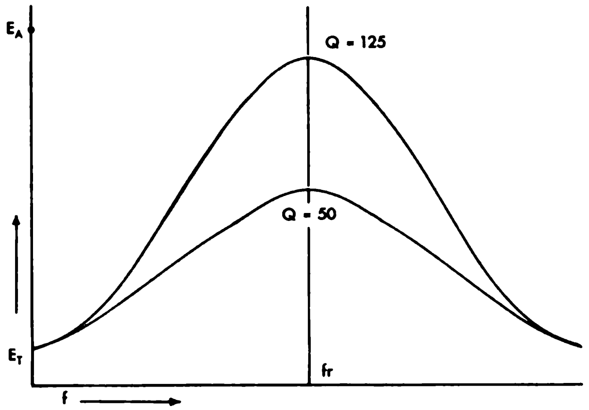


Parallel LC Circuit

Specifically, the greater the R_G of the circuit, the greater the selectivity. In contrast, the *smaller* the R_L of the circuit (or the greater the Q of the coil — $Q = X_L/R_L$), the greater the selectivity.

The following illustrations show these two principles graphically.





Note the relative sharpness of the curves under the different conditions. Remember, the steeper the slope of the curve, the more narrow the bandwidth and the better the selectivity.

In summary, it can be stated that to secure high selectivity with a *series* LC circuit, high Q coils and generators with *low* internal resistance should be used. To secure high selectivity, with *parallel* LC circuits, high Q coils and generators with *high* internal resistance should be used. Since most generators in radio and radar circuits are vacuum tubes, which have high internal resistance, parallel LC circuits are used much more extensively than series LC circuits.

Series and Parallel Resonance Compared

Study the table below which summarizes the important facts about series and parallel resonant circuits.

Series Resonance	Parallel Resonance
$Fr = \frac{1}{2\pi\sqrt{LC}}$	$Fr = \frac{1}{2\pi\sqrt{LC}}$
$X_L = X_C$	$X_L = X_C$
PF = 1	PF = 1
Z is minimum	Z is maximum
I is maximum	I is minimum
E is minimum	E is maximum
Bandwidth = .707 I_{max}	Bandwidth = .707 E_{max}

APPLICATIONS OF RESONANT CIRCUITS

There are two main uses for parallel resonant circuits in communications equipment. The most important of these is the selection of a single frequency or a narrow band of frequencies. This is called *tuning.* This is possible since the impedance across a parallel resonant circuit and consequently the voltage, is maximum at the resonant frequency and decreases to very low values on either side of resonance. Some circuits of this type are made up of fixed components so that the circuit has a fixed frequency. The most common, however, is the variable-tuned circuit. This type of circuit can be tuned to a certain frequency or narrow band of frequencies within a wide band of frequencies. Tuning is usually accomplished by a variable capacitor. The minimum amount of capacity and the fixed inductance determine the highest frequency that can be selected, and the maximum amount of capacity and the fixed inductance determine the lowest frequency. A common variable capacitor used in commercial broadcast receivers has a minimum capacity of about 35 mmf and a maximum capacity of about 365 mmf. When this capacitor is connected in parallel with a .28 millihenry coil, the circuit can be tuned to any frequency in the broadcast frequency range (550-1600 kc). Some circuits consist of a fixed capacitor and a variable inductor. Variable inductors are usually of the powdered-iron core type as mentioned on page 98.

The second application of a parallel-resonant circuit is as a frequency rejection *filter* or trap. When used for this purpose, the circuit is placed in series with a line delivering voltages at the desired frequencies to the proper circuit. Since the impedance across this type of circuit is maximum at resonance, it will drop most of the voltage at the unwanted frequency and pass most of the voltages at the desired frequencies.

A parallel resonant circuit that is used for tuning may also be called a series resonant circuit. Within itself, it is a series resonant circuit. When a voltage is induced into the coil part of the circuit, the coil actually acts as the source of current. Therefore, this effective voltage source is in series with the inductance of the coil and with the capacitor. Maximum current flows within the circuit. This maximum flow characteristic is the distinguishing feature of a series resonant circuit. An application of a series resonant circuit is, therefore, also that of frequency selection or tuning. Series resonant circuits are also used to reject one frequency or a narrow band of frequencies from a wide band of frequencies.

Chapter 19

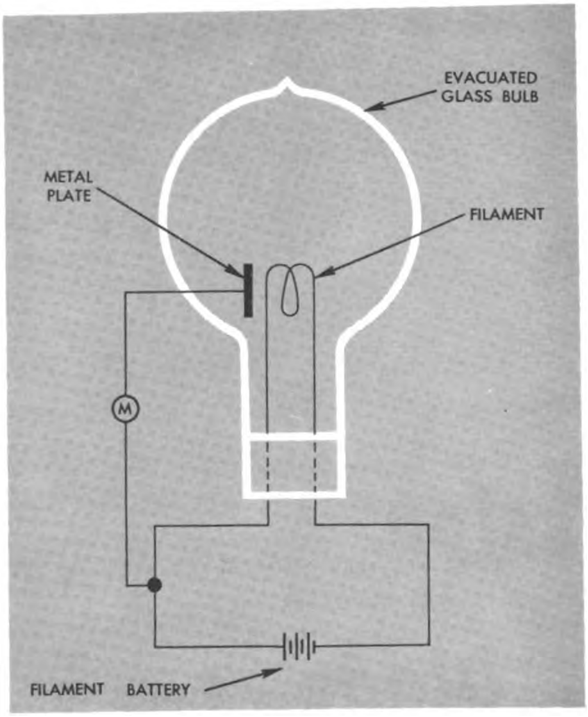
Electron Emission and the Diode

Electronics has been defined as the science dealing with the control of electron flow, especially by means of electron (or vacuum) tubes. The terms *electron tube* and *vacuum tube* are used interchangeably to refer to a tube in which a flow of current takes place in an enclosed space. A vacuum tube is one from which the air has been removed to form a partial vacuum. However, in some instances certain gases are introduced inside the tube. Although the enclosed space is not a vacuum, these gas tubes are included in the category of vacuum or electron tubes.

There are numerous types of vacuum tubes manufactured; each type differs from the other and has specific characteristics. In this chapter, some theories, principles, and facts that apply to vacuum tubes in general are presented. The various types are then discussed separately in the remainder of this chapter and those which follow.

EDISON EFFECT

In 1883, the electric light bulb, invented by Thomas A. Edison, was a very faulty device. The lamp consisted of a glass bulb, inside of which a thin conductor, the filament, was secured. The air was then pumped out of the bulb and the bulb sealed. When direct current flowed through the filament, the filament was heated to incandescence. The amount of light given off by the filament was insignificant by modern standards; in addition, the filament lasted a comparatively short time. In an attempt to improve the original electric light bulb, Edison and his associates conducted experiments, using various materials as filaments. During these experiments, they noticed that a dark deposit formed on the inside of the glass, reducing the amount of light radiated. In an unsuccessful attempt to eliminate this deposit, a metallic plate was introduced inside the glass



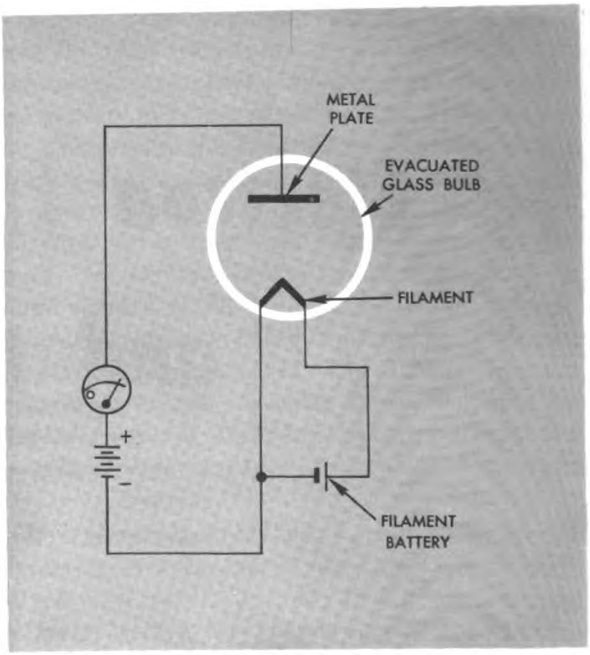
Edison's Experiment

bulb. A sensitive meter was connected between the metal plate and the filament of the lamp as shown in the illustration above.

Edison was amazed to discover that a flow of current was indicated by the meter. He recorded his findings in his notebook, but it was left to others who followed him to make use of his discovery ,to develop the new science of electronics. In his honor, the emission of electrons by a heated filament is called the *Edison effect.*

FLEMING'S VALVE

During the twenty years that followed the discovery of the Edison effect, the electron theory of current flow was developed. In 1904, John Ambrose Fleming, an Englishman, explained the Edison effect, using the electron theory. Fleming constructed a circuit similar to Edison's to substantiate the assumption that the heated filament was emitting negatively charged particles (electrons). He reasoned that if the emitted particles were negative, a positive charge should attract them. He proved his point by noting that current flowed when the plate was positive with respect to the filament, but no current flowed through the meter when the battery between the plate and the filament of this circuit was reversed. See illustration, Fleming's Experiment. This result is exactly what is expected on the basis of the electron theory. When the plate is positive, electrons are attracted to it; when the plate is negative with respect to the filament, the emitted electrons are repelled.

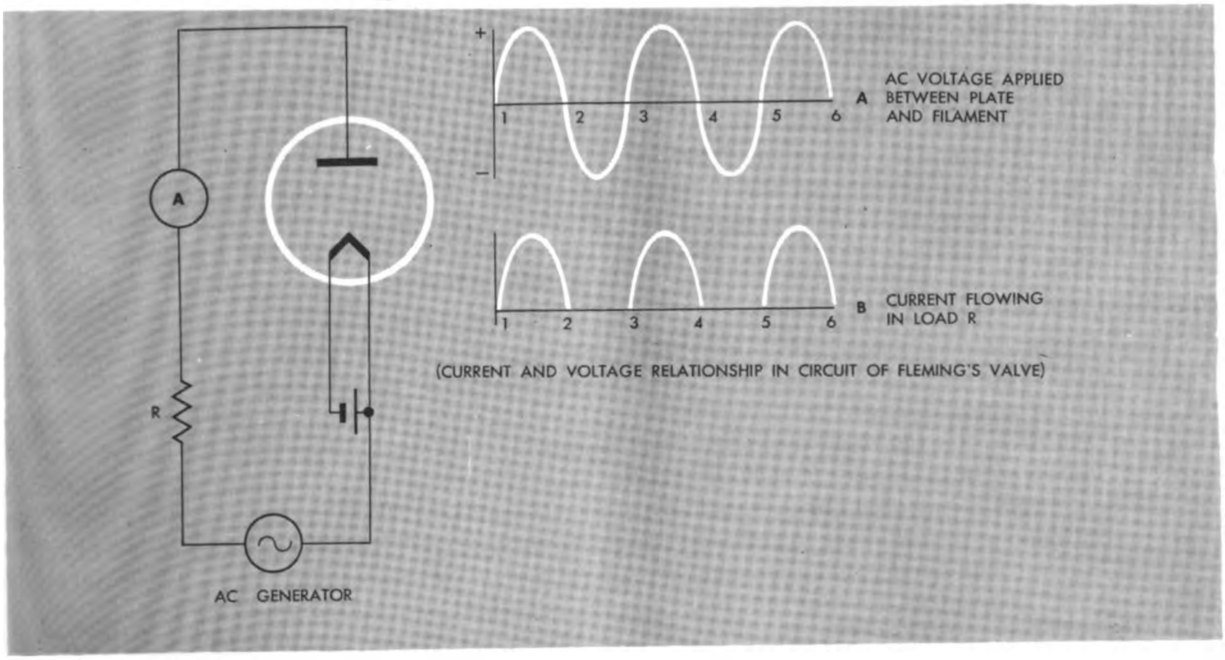


Fleming's Experiment

The next step was to replace the battery between the plate and the filament by an AC generator as shown in the illustration on page 151. Fleming found that, although he used an AC generator, the current which flowed was a direct current. This is explained as follows:

Current flows during the portion of the AC cycle when the plate is positive with respect to the filament; during the portion of the cycle when the plate is negative, no current flows. As a result, the current flow is a pulsating direct current as shown.

Since the device acts like a valve, permitting current to flow in only one direction, Fleming called it a *valve*. Thus, the first vacuum tube was known as Fleming's valve. The English still refer to vacuum tubes as valves.


A
R
AC GENERATOR
+
−
1
2
3
4
5
6
A
AC VOLTAGE APPLIED BETWEEN PLATE AND FILAMENT
B
CURRENT FLOWING IN LOAD R
(CURRENT AND VOLTAGE RELATIONSHIP IN CIRCUIT OF FLEMING'S VALVE)

Fleming's Valve

ELECTRON EMISSION

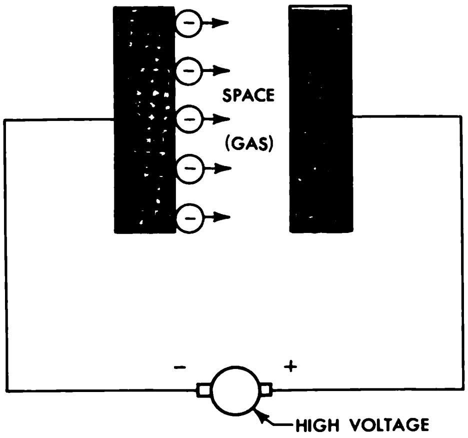
All vacuum tubes have one characteristic in common: free electrons are forced to leave the surface of a conductor and enter the space around it. *Electron emission* is the term applied to electrons leaving the surface of a solid. For electrons to be emitted, energy must be applied in some form. The emission of electrons may be accomplished by any one of these methods: *field emission, secondary emission, photoelectric emission,* and *thermionic emission.*

Field Emission

The field emission of electrons may be obtained by subjecting a substance to an intense electrostatic force. If a high voltage is connected between two substances A and B so that A is negative and B is positive as shown to the right, the free electrons in A experience a force attracting them to B. If the force attracting the free electrons is greater than the binding force of the atoms and molecules, electrons leave the surface of A and enter the space between A and B. The amount of emission obtained in this manner depends on the value of the voltage applied, the distance between the positive and negative poles, and the number of free electrons in the substance. Some evidence of field emission you have observed are electric sparks, lightning, and the breakdown of electrical insulation.

Secondary Emission

Secondary emission of electrons occurs if a substance is bombarded by atomic particles moving at high velocities. These particles may be positively charged, negatively charged, or neutral. If the electrons emitted from A in the illustration, Field Emission, are acceler-



Field Emission

ated so that they strike B with sufficient force, free electrons must be dislodged from the surface of B. These electrons are then emitted into the space between A and B by secondary emission. The amount of secondary emission depends upon the number and velocity of the moving electrons and the number of free electrons in the substance they strike.

Photoelectric Emission

The photoelectric emission of electrons occurs if a substance emits electrons when light hits it. The amount of emission depends upon the intensity of the light, the wavelength of the light, and the type of substance used. You have seen the effects of photoelectric emission in the following applications: a pickup tube for television transmission (an iconoscope), photographic exposure meters, and automatic door openers.

Thermionic Emission

The thermionic emission of electrons occurs if a substance emits electrons when heat is applied to it. Metallic substances conduct electricity as a result of the motion of free electrons. At ordinary temperatures these free electrons cannot escape from the surface of the metal because of attracting forces at the surface. The attracting forces make up what is called the *surface barrier*. As you heat the metal, the velocity of the free electrons increases and they acquire more energy. When the electrons acquire enough energy to overcome the surface barrier, they escape. The substance from which the electrons escape is called the emitter or *cathode*.

Note, in the illustration to the right, that the metal which has lost electrons has become positive and attracts the free electrons back to its surface. The space immediately surrounding the metal is negative because of the emitted electrons and repels free electrons. Thus electrons are continually leaving and returning to the surface of the heated metal. The result is the formation of a *cloud* of electrons around the emitter.

The number of electrons emitted by thermionic emission depends on the amount of heat and the substance to which the heat is applied. The emission of electrons in vacuum tubes is accomplished almost exclusively by thermionic means.

EMITTERS

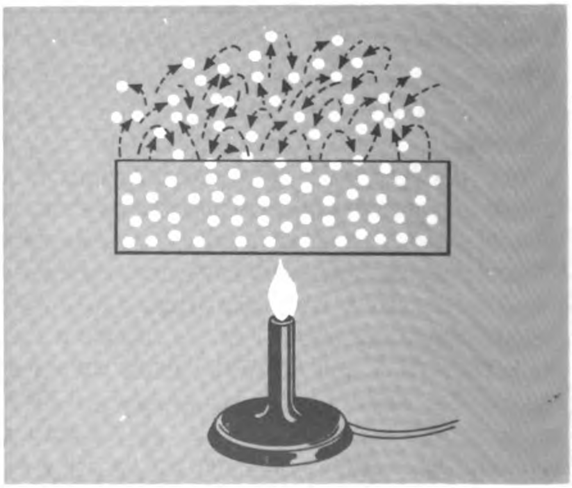
An emitter, or cathode, to be used in a vacuum tube must be capable of emitting a large quantity of electrons when heated. It must also have a long life and not be damaged too easily. Therefore the material of which it is constructed must be carefully chosen.

Emitter Materials

The number of materials suitable for thermionic emitters is limited. A good thermionic emitter must have a low *work function*. The work function can be defined as a measure of the amount of energy required to release an electron. The emitter must have good mechanical properties at high temperature; that is, it must be able to withstand physical movement or jar. Its melting point must be higher than the temperature necessary for emission.

Only two pure metals fulfill the above conditions, *tantalum* and *tungsten*. Tantalum has a lower vaporization temperature than tungsten and is therefore rarely used. Tungsten requires a very high temperature for proper emission. It has, however, great durability and is used in large, high-voltage tubes (above 3500 volts), where other types of emitters would not last long.

In 1913, Langmuir discovered that if thorium oxide were added to the tungsten powder before the filament was made, the result was a filament material of greatly improved emission



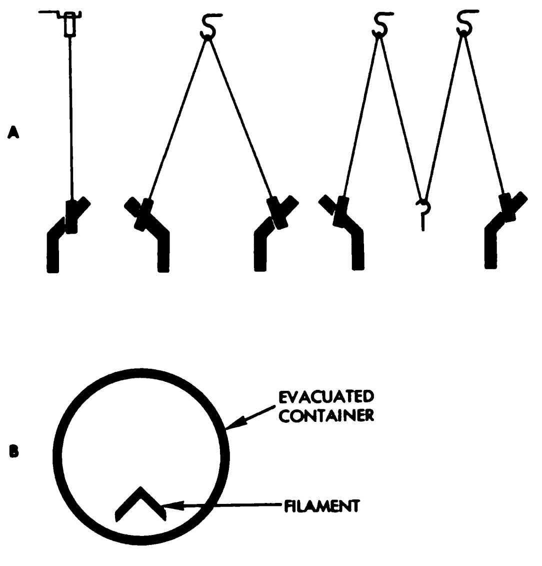
Thermionic Emission

characteristics. When this material is heated to a high temperature, some of the thorium oxide reduces to metallic thorium, diffuses to the filament surface, and coats the tungsten with a layer of thorium one molecule thick. Thorium is positive with respect to tungsten; so an intense electric field which aids the electrons to escape is set up at the tungsten surface. The addition of thorium therefore reduces the work function so that the operation of the filament at normal temperature causes the emission of many electrons. Gradual evaporation of the thorium layer may be balanced by additional diffusion of thorium to the surface, if proper operating temperatures are maintained. If operated at either too high or too low a voltage, a loss of emission may result; but if operated at proper temperature, the life of a tube is limited only by the reserve supply of thorium stored in the tungsten. Cathodes employing this combination of thorium and tungsten are known as *thoriated-tungsten cathodes.*

Wehnelt discovered, as early as 1904, that emission from the oxides of the rare earth metals was large even at low temperatures. Not much use was made of this knowledge until the early 1920's when oxide-coated filaments were introduced to reduce filament power consumption of the battery-operated radios. The oxide-coated cathode consists of a metal base—nickel, konel, or platinum alloy — coated with a layer of alkaline oxide, usually barium or strontium oxide. It is believed that in the activation process, some of the oxide breaks down and a one-molecule layer of barium or strontium is deposited on the surface. The surface layer of barium or strontium is positive with respect to the oxide. The source of electrons is believed to be the oxide whose work function is reduced by the positive layer of metal. Oxide-coated emitters have a long life and great emission efficiency. They are used satisfactorily at operating temperatures of from 800-900° C. This temperature is low compared to approximately 1650° C for thoriated tungsten and 2200-2350° for pure tungsten.

Types of Emitters

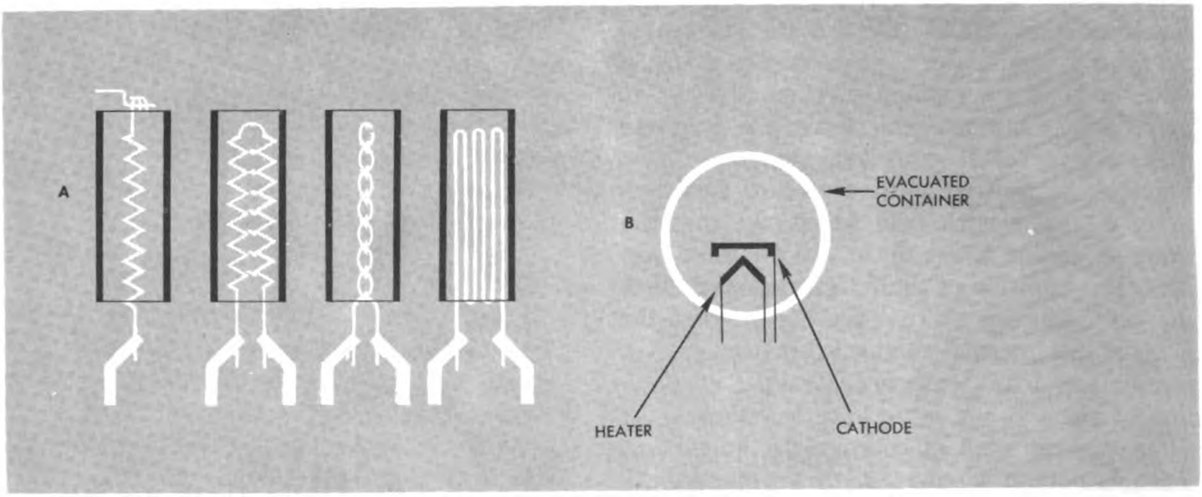
The electron-emitting cathodes of vacuum tubes are divided into two general classes: (1) directly heated (or filamentary) cathodes and (2) indirectly heated (or equipotential) cathodes.



A. Directly Heated Electrodes
B. Schematic Representation

In directly heated cathodes, as shown above, a wire is heated by the passage of electric current through it, and electron emission takes place directly from the wire. Since the heating current flows through the wire, a potential drop of a few volts is distributed along the wire; that is, all points of the filament are not at the same potential. The emission from all points on the cathode is therefore not uniform. When AC is used to heat the filaments, the stray alternating electrostatic field, and the alternating voltage across the filament, may bring about undesirable fluctuations of plate current.

The difficulties encountered in using directly heated cathodes led to the development of the indirectly heated cathode. As shown on page 154 a twisted hairpin or loop of tungsten wire, insulated with a baked on material, is inserted in a thin-walled hollow metal cylinder, usually made of nickel. The outside of the cylinder is coated with material which emits electrons. Passage of current through the tungsten heater causes the cylinder to become hot enough to emit the desired electrons. Since the emitted current is usually quite small and no heating current flows in the cylinder, there is no voltage drop along the cylinder. The cathode is *equipotential;* that is, all points on the cathode are at the same potential.

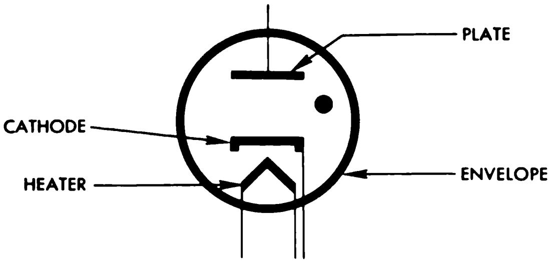


A. Indirectly Heated Cathode B. Schematic Representation

Tungsten, thoriated tungsten, and oxide-coated emitters may be of the filamentary or directly heated type, but indirectly heated cathodes use oxide-coated emitters. The heater of an indirectly heated cathode is not part of the vacuum tube circuit. Its only function is to heat the cathode.

Space Charge

When an emitter is heated to the required temperature, electrons gain sufficient energy to be liberated from the cathode. If these emitted electrons are not drawn away by an external field, they will form a *space charge*. This was described on page 152 as a cloud of electrons which are moving about with different velocities. Since electrons have like charges, they repel each other; some of them tend to drift away from the emitting surface. Collisions between electrons cause some of them to head back toward the emitter with enough speed to re-enter it. The accumulation of the negative electrons near the emitter repels others coming from the emitter. The emitter, if insulated, becomes positive because of the loss of electrons. An electrostatic field is therefore, established between the cloud of negative electrons and the now positive cathode. A balance is finally reached when only enough electrons move from the cathode to the area surrounding it to supply the loss caused by diffusion of the space charge.



Schematic Symbol for a Gas Diode

DIODE TUBES

Every vacuum tube must have at least two elements or electrodes: a *cathode* and an *anode* (plate). The plate is the electrode toward which the electrons emitted by the cathode are attracted. The material, shape, size, and location of the plate vary with the type of tube. A typical plate consists of a conductor which completely surrounds the cathode.

In addition to the cathode and plate, other electrodes — which are discussed later — may be introduced in the tube. The number of electrodes serves as one method of classifying tubes:

A 2-electrode vacuum tube is a *diode*.
A 3-electrode vacuum tube is a *triode*.
A 4-electrode vacuum tube is a *tetrode*.
A 5-electrode vacuum tube is a *pentode*.
Fleming's valve was a diode vacuum tube.

Enclosing Tube Elements

If an emitter were heated to the required temperature in air, it would oxidize and be destroyed in a short time. For this and other reasons, it is essential to enclose the elements of a vacuum tube in a space as free of air as possible. There are two classes of vacuum tubes: ***high-vacuum*** and ***gas tubes.***

The ***high-vacuum*** or ***hard*** tubes have their elements inclosed in a space where the pressure has been reduced to about 10.3 microns. A micron of gas pressure is defined as that pressure which will support a column of mercury .0001 centimeter high. To appreciate the degree of vacuum obtained in electron tubes, compare these figures to normal sea level pressure of about 760,000 microns.

The ***gas*** or ***soft*** tubes have their elements in a space from which the air has been evacuated and a specific gas has been admitted at pressures ranging from one to five hundred microns.

The glass or metal enclosing the high vacuum or gas in electron tubes is called the ***envelope.*** It is indicated in schematic drawings as a circle or an ellipse. To distinguish between a vacuum and a gas tube on drawings, a dot is placed inside the circle (or ellipse) of the symbol for gas tubes. The schematic symbol for a gas diode are shown in the illustration at the bottom of page 154.

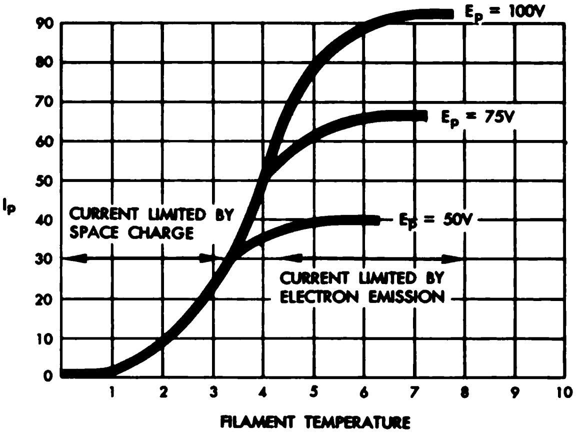
In the manufacture of high-vacuum tubes, many precautions are taken to insure the removal of gas from walls and electrodes. The electrodes are thoroughly cleaned and are then heated for several minutes in an atmosphere of hydrogen, which removes oxygen and water vapor. After the tube is assembled and connected to the pumps, the electrodes are heated by high frequency induction in order to remove other gases. Residual gas is removed by the use of a ***getter*** — an active chemical substance such as barium, magnesium, aluminum, or tantalum — which has the property of combining with gases when it is vaporized. In glass tubes, a small amount of the getter is mounted in such a position that it will be heated and vaporized during the inductive heating of the elements. The result of this may be seen as the dark spot on the glass near the base of most glass tubes.

Current Flow in a Diode

Experiments with diode vacuum tubes reveal that the amount of current which flows from the cathode to the plate depends upon two factors: the ***temperature*** of the cathode and the ***potential*** between the cathode and the plate.

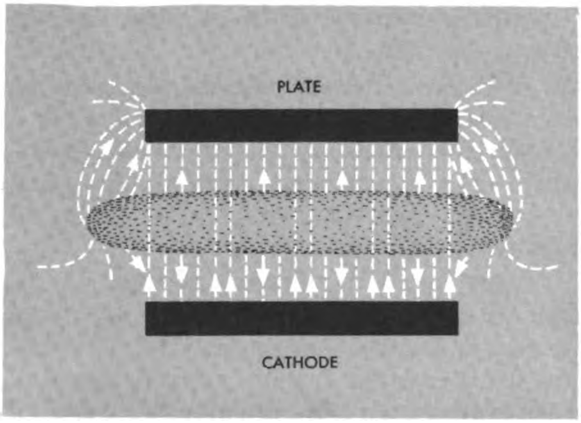
The distance between the cathode and the plate also affects the current which flows in a vacuum tube. Since the distance between elements in a vacuum tube is fixed, this effect is of no practical value except to designers and manufacturers of vacuum tubes.

EFFECT OF CATHODE TEMPERATURE. As the temperature of the cathode is increased by an increase in filament current (with the plate voltage kept constant), plate current increases. This increase in plate current continues until a point is reached where no further increase occurs regardless of how much the filament temperature is raised. This condition — plate current ceases to increase appreciably with an increase in cathode temperature — is called ***temperature saturation*** or ***space charge limitation.*** This effect for different values of plate voltage is shown in the illustration below. Note that for each value of plate voltage a point is reached where an increase in temperature results in no appreciable change in current.



Space Charge Limitation

The effect of space charge in limiting emission may be understood by analyzing the electrostatic fields within the diode. As shown in the next illustration, two electrostatic fields exist in the diode — one between the cathode and plate, and the other between the cathode and the space charge. The force exerted on the emitted electrons is due to the difference between these fields which are of opposite polarity. At low cathode



Electrostatic Fields within the Diode

temperatures, the plate attracts almost all the electrons emitted. However, when a temperature is reached at which the electrostatic field of the space charge becomes large enough to nullify the attracting field of the plate, additional current ceases to flow to the plate.

EFFECT OF PLATE VOLTAGE. In practical applications of vacuum tubes, the voltage applied to the filament or heater is a fixed value. This value is usually one recommended by the manufacturer to give maximum emission and long life to the cathode of the particular tube. Therefore, a knowledge of the effects of plate voltage on plate current is of more practical value.

As the plate voltage of a diode is increased while the filament voltage is kept constant, plate current increases until a point is reached where all the electrons emitted by the cathode are attracted to the plate. A further increase in plate current is no longer possible regardless of plate voltage increase. This condition — plate current ceases to increase appreciably with an increase in plate voltage — is called *plate saturation* or *voltage saturation*.

Note in the illustration of Plate Saturation that a point is reached where an increase in plate voltage results in no appreciable increase in plate current. The slight increase in plate current that occurs as the plate voltage is increased beyond saturation is due to the increased electrostatic field between plate and cathode. This increased electrostatic field has the effect of increasing the emission from the cathode; it is known as the *Schottky effect*.

Characteristic Curves

Curves similar to the one illustrated to the right may be plotted for any diode. These curves, in which values of I_p are plotted for increasing values of E_p (with I_f constant), are called E_p-I_p, or characteristic, curves. The values vary for different diodes. The characteristic curve of a diode will help determine to a large extent how you can use the tube.

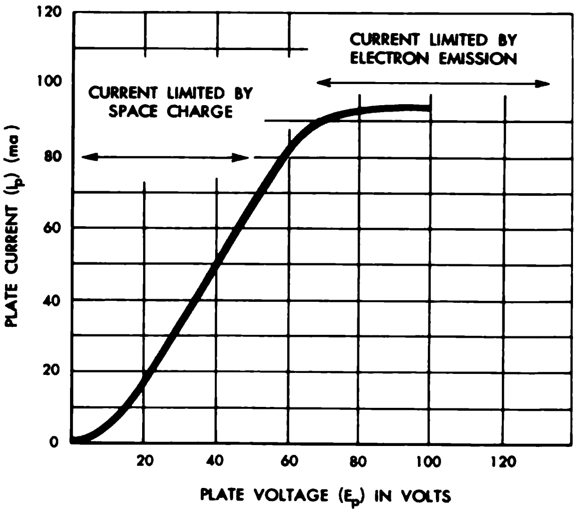
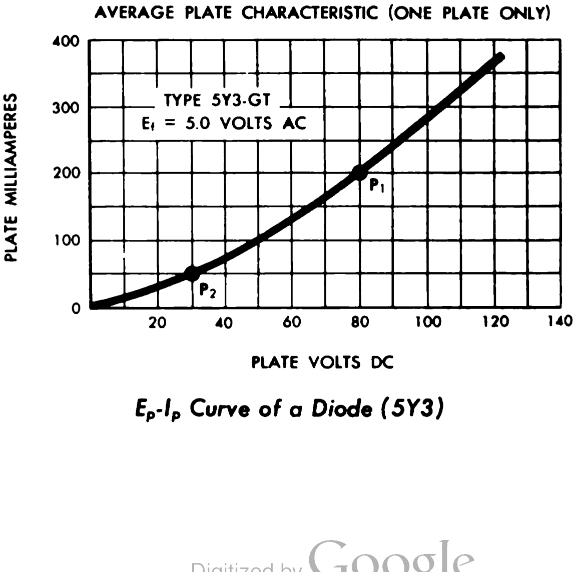
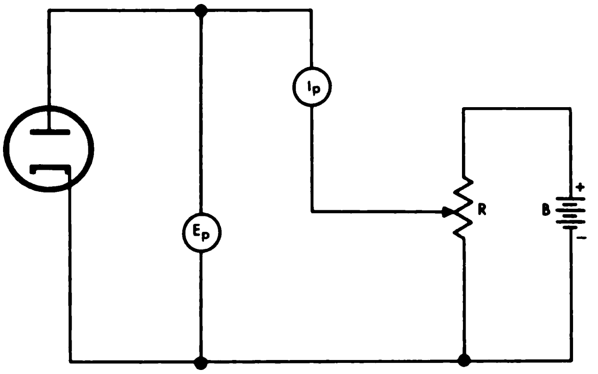


Plate Saturation

You may plot a characteristic curve for any given diode by using the circuit shown in the illustration on page 157. Vary the plate voltage and read the value of plate current obtained for each value of plate voltage. From these values you may draw a curve similar to the one shown by plotting the plate voltage along the horizontal axis and the plate current along the vertical axis.



E_p-I_p Curve of a Diode (5Y3)



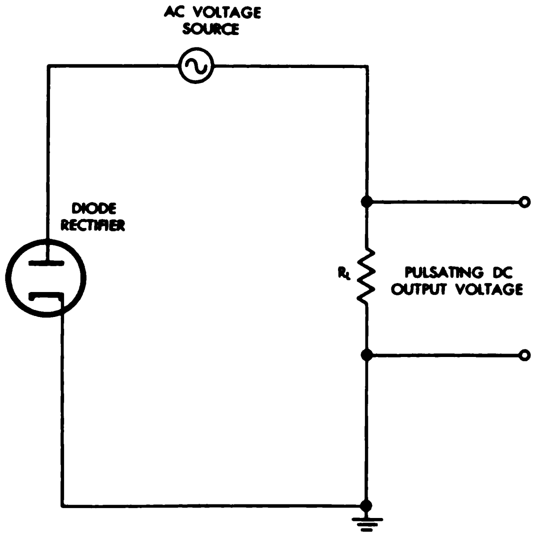
Circuit for Plotting an E_p-I_p Curve

The E_p-I_p curve of a given diode is useful because it yields the following information rapidly:

The value of plate current for any value of plate voltage applied.

The value of plate voltage needed to obtain a given value of plate current.

In addition, DC plate resistance may be obtained from the E_p-I_p curve of a diode. DC plate resistance may be defined as the opposition the tube offers to plate current when the plate voltage is kept constant. Using the E_p-I_p curve of the 5Y3 shown at the left and applying Ohm's law, you find that the DC plate resistance depends upon which point of the curve is chosen.



A Simple Rectifier Circuit

At point P_1, $E_p = 80$ volts, and $I_p = 200$ ma. Therefore,

$$R_p = \frac{E_p}{I_p}$$

$$R_p = \frac{80}{.2} = 400 \text{ ohms}$$

At point P_2, $E_p = 30$ volts, and $I_p = 50$ ma. Therefore,

$$R_p = \frac{30}{.05} = 600 \text{ ohms.}$$

RECTIFICATION

A diode conducts only when its plate is positive with respect to the cathode; as a result, an AC signal applied between plate and cathode yields a direct current. A diode may therefore be used as a *rectifier*.

A Simple Rectifier

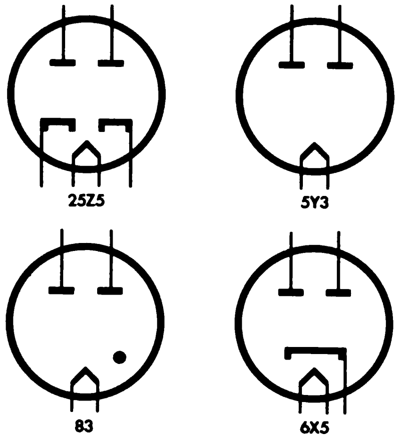
The schematic of a simple rectifier circuit consists of a diode and load resistor in series as shown. On the positive alternation, the plate of the diode is positive with respect to the cathode, and the tube conducts. The current flowing through the resistor causes a voltage drop. On the negative alternation the plate is negative with respect to the cathode, and the tube does not conduct. There is no voltage drop across the resistor during this alternation. The output voltage is therefore a pulsating DC.

Inverse Peak Voltage

A diode used as a rectifier will not conduct if the plate is negative with respect to the cathode. However, the tube will arc over if too great a voltage of this polarity is applied across the tube. The voltage that a diode will stand (in the opposite direction) without arcing is called its *inverse peak* voltage. This value varies for different tubes and must be considered in selecting a tube. Manufacturers list the inverse peak voltage of each diode in their tube manuals. For example, the inverse peak voltage of a 6H6 is listed as 420 volts. This means that if the plate is made 420 volts negative with respect to the cathode, the tube may be damaged.

Types of Rectifiers

Rectifiers may be divided into two general types. In one type, the rectified current is large and is used for power purposes. In the second type, the rectified currents are small and are

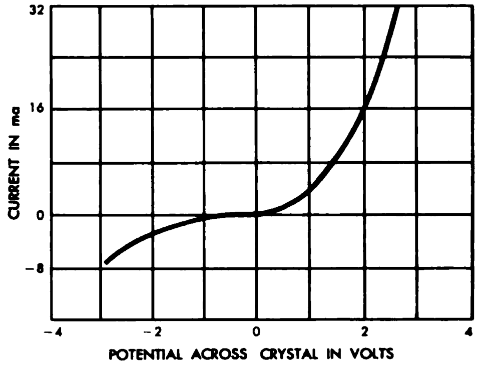


Schematic Representation of Twin Diodes

used for signal or control purposes. The plates, cathodes, filaments, and element spacing of diodes used for power rectification are large and ruggedly constructed in order to withstand higher values of current and voltage. In many instances, gas-filled diodes are used in preference to vacuum-tube diodes because the gas tube permits the use of greater values of current. Since the values of current and voltage encountered in signal-type diodes used for detection are small, the elements and spacings are correspondingly small.

The simple diode tube has many applications which you will encounter. You will find that many circuits require the use of two diodes. As a result of the need for two diodes, and in the interest of economy, tube manufacturers make vacuum tubes containing two diodes in a single envelope. Several such tubes, known as duodiodes, or twin diodes, are shown schematically above. Note that both directly and indirectly heated cathodes are used; also, that twin diodes must have two plates but may have either one or two cathodes.

Certain crystalline materials, such as silicon, carborundum, galena, and germanium have electrical characteristics which classify them as semi-conductors. These materials exhibit greater resistance to current flowing in one direction than to current flowing in the opposite direction. The ratio of these resistances (the front-to-back ratio) is usually 10:1 or greater. A characteristic curve of such a crystal is shown below.



Characteristic Curve of a Crystal

Note that the portion of the curve to the right of the 0-volt line is very similar to the E_p-I_p curves of vacuum tube diodes. However, unlike the vacuum-tube diode, a small amount of current does flow when the polarity of the voltage is reversed. In many applications, the flow of the small value of current is of little consequence; this permits the use of such crystals as rectifiers. Because of their similarity to the vacuum-tube diode, crystals are frequently referred to as *crystal diodes*. Crystals cannot be used in circuits having large values of current and voltage. They are used primarily as detectors of very high radio frequency signals.

Power supplies, principles of rectification, and filtering are discussed more fully in chapter 25.

Chapter 20

Triodes

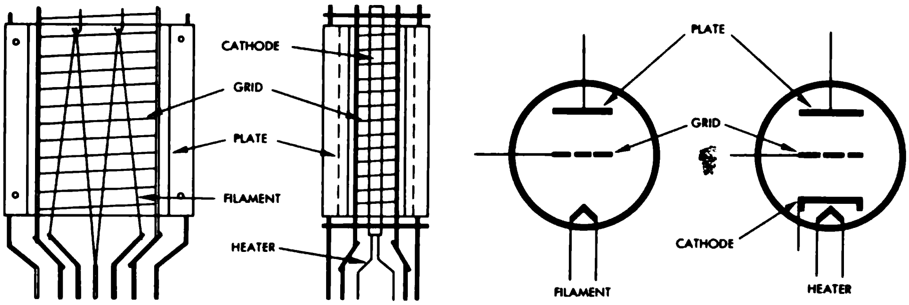
The remarkable achievements in the field of electronics — which includes radio, radar, television, guided missiles, etc. — are a direct result of the discovery of the vacuum tube. The greatest single advance in the development of vacuum tube theory and application was the audion tube of Lee DeForest. The outstanding feature of this tube was the insertion of a third electrode in the diode that Fleming had developed. DeForest used this third electrode which was placed between the cathode and the plate of the diode to control the amount of electron flow. Because of the form of this control electrode in the earlier tubes, the term *grid* was applied to it. The grid is now defined as an electrode that contains openings through which electrons may pass. Physically, it is constructed of mesh, screen, or spiral wires placed between the cathode and the plate. The construction and schematic representation of the grid are shown in the illustration on the next page.

Many modern tubes contain more than one grid; so, to describe its function as well as to differentiate it from other grids, this third electrode is called a *control grid*. A tube which contains three active elements — a cathode, a control grid, and a plate — is called a *triode*.

EFFECT OF CONTROL GRID

Since it is located between the plate and cathode, the control grid is able to control the plate current of a triode by variations in its potential. How this control is exerted by means of the electrostatic fields set up between the cathode, and the other two electrodes can be best understood by examining the illustration at the bottom of the next page.

An electrostatic field is conventionally represented with arrows pointing from the more positive to the more negative potential — in the direction a unit positive charge moves if placed in the field. Since an emitted electron is a unit



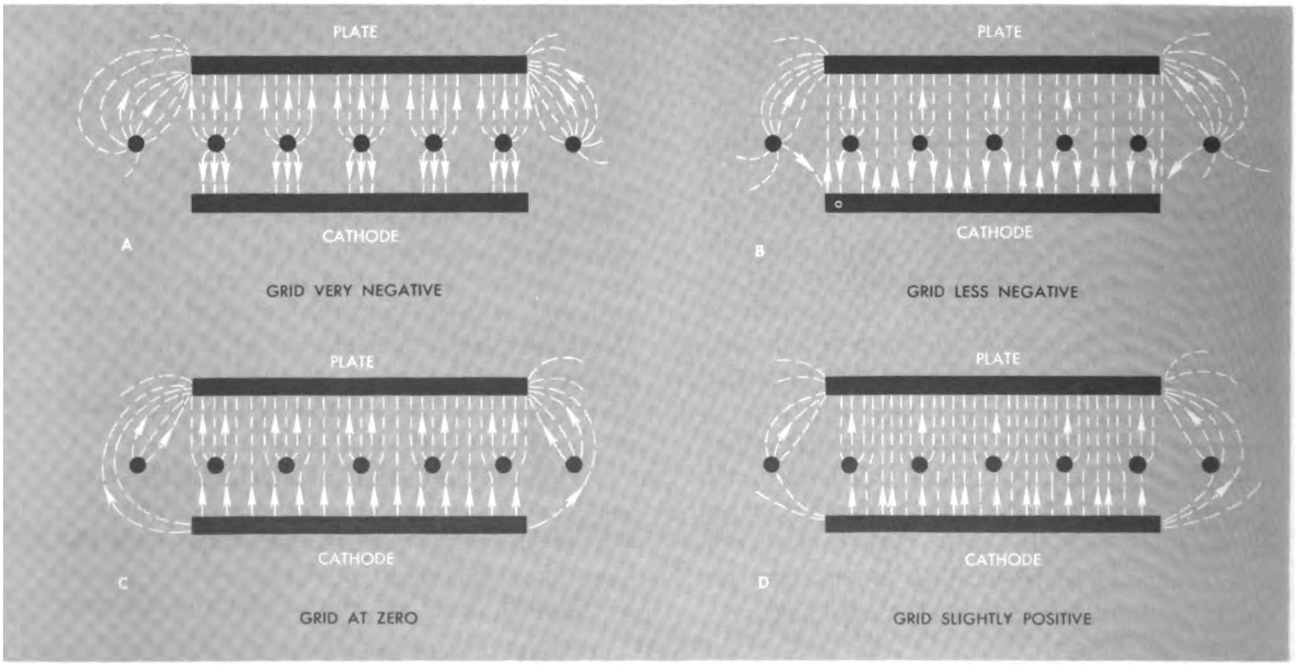
Construction and Schematic Representation of a Triode

negative charge, the direction the electron moves in an electrostatic field is the reverse of this conventional placement of arrows. In the illustration, the electrostatic field arrows are drawn to indicate the direction in which the *emitted electrons are urged.*

In diagram A, the control grid is made so negative with respect to the cathode that all lines of force terminate on the grid and no field exists between the cathode and plate. This field is in such direction as to return electrons emitted from the cathode back toward it. Thus, there are no electrons that reach the plate, and the plate current is zero.

If the grid is made less negative, as shown in diagram B, the electrostatic fields consist of two forces: first, a force (plate potential) which urges electrons from the cathode directly to the plate, and second, a force (grid potential) which repels the electrons emitted from the cathode back to it. In this case, plate current flows and increases to the extent that the first force is greater than the second. Thus, as the grid is made less and less negative, the plate current flow increases. Since the initial velocity of the emitted electrons is low, a relatively small negative voltage on the grid repels these emitted electrons away from the grid and very little (or no) grid current flows. As this negative voltage on the grid is further decreased (below 1 volt) and finally becomes zero (diagram C), there is only one force (plate voltage) which affects

Electrostatic Fields in a Triode



the electrons. Plate current is high under these conditions, since the emitted electrons no longer experience a retarding force. Some electrons strike the control grid and cause a small grid current flow.

As the control grid is made more positive than the cathode (diagram D), both the grid and plate potentials cause electrostatic fields to be set up which attract the emitted electrons. Both plate current and grid current flow are considerable under these conditions. *Space current* is the term applied to the *total current* flowing from the cathode; in the case of the triode, it is the sum of the plate current and the grid current.

AMPLIFICATION FACTOR

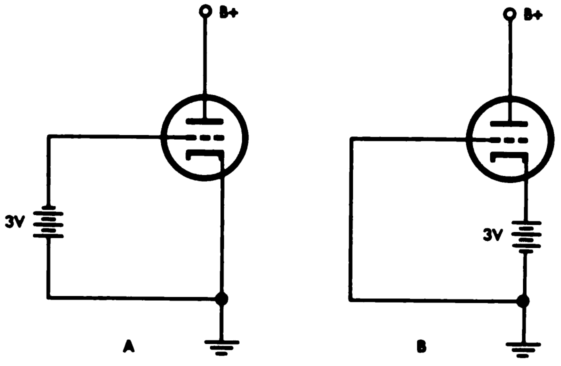
The preceding discussion of the electrostatic fields in a triode showed that the net electrostatic field at the cathode is the resultant of three fields. These fields are produced by: the potential on the grid (E_g), the potential on the plate (E_p), and the space charge. An increase in plate current occurs if the *grid voltage is made less negative* or if the *plate voltage is made more positive.*

Because the grid is physically closer to the cathode and also screens a portion of the field extending from the plate to the cathode, a grid voltage change affects plate current much more than a plate voltage change. Thus, a change of one volt in grid voltage produces much greater variation in plate current than a change of one volt in plate voltage. To put it another way, it requires a larger voltage change on the plate than on the grid to cause the plate current to vary a given amount. The ratio of the change in plate voltage to the change in grid voltage required to cause a given plate current change is called the amplification factor of the tube. It is represented by the Greek letter μ (mu) and is always greater than unity. For example, if the same change in plate current is produced by a change of one volt on the grid or of ten volts on the plate, the grid is ten times as effective as the plate in controlling plate current; therefore, μ is 10.

BIAS

You have seen that the grid voltage has a great effect on the flow of plate current in a triode. By varying this voltage from a very negative to a slightly positive voltage, plate current flow varies from zero to a maximum. Thus the grid may be considered to be a valve, and variations in its potential cause either an increase or decrease in the flow of electrons to the plate. This fact is made use of in the operation of triodes.

The term grid voltage as used here means the voltage of the control grid with respect to the cathode. In most applications the grid is usually at a negative potential with respect to the cathode. This negative potential of grid with respect to the cathode is called *bias*. If you say that the bias is 3 volts, you mean that the grid is 3 volts *negative* with respect to the cathode. Two methods of obtaining a bias of 3 volts are shown. In diagram A the battery is connected with its negative terminal to the grid, while its positive terminal and the cathode are grounded. In diagram B the positive terminal of the battery is connected to the cathode, while its negative terminal and the grid are grounded. In both cases the grid is 3 volts negative with respect to the cathode. If the grid and cathode are at the same potential, there is no difference in potential and the tube is operating with zero bias. Replace the battery in either diagram A or B by a conductor, and the tube is at zero bias.



Obtaining a Bias of Three Volts

Remember. Bias is a *DC voltage* applied between the grid and cathode of a tube, and increasing the bias means *making the grid more negative*. If bias is made so great that no plate current flows, the tube is at *cutoff* bias. The bias at which a triode will reach cutoff depends upon the tube and the plate voltage applied. In early days of radio, bias was provided by a battery called a C battery; a common symbol for bias voltage is therefore E_c. However, obtaining bias by means of a battery, as shown in the illustration, is only one of many methods now used. Other methods are discussed later in this chapter.

In addition to the DC bias, a DC voltage may be applied directly between the grid of a tube and ground. In such cases, the voltage between the grid and ground is represented by the symbol E_g.

CHARACTERISTIC CURVES

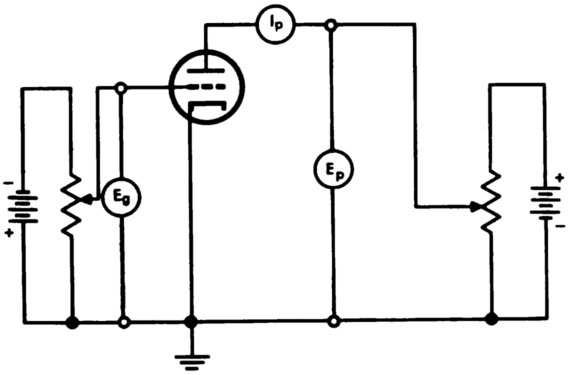
The relationships existing between grid voltage, plate voltage, and plate current are frequently expressed graphically by characteristic curves. Two sets of these characteristic curves are used very frequently in triode applications, since they may be used to predict, to some extent, the behavior of a given triode when certain electrode voltages are applied. These are known as:

Mutual characteristic (E_g-I_p) *curves* — obtained by plotting the plate current (I_p) against the grid voltage (E_g) while keeping the plate voltage (E_p) constant.

Plate characteristic (E_p-I_p) *curves* — obtained by plotting the plate current (I_p) against the plate voltage (E_p) while keeping the grid voltage (E_g) constant.

Mutual Characteristic Curves

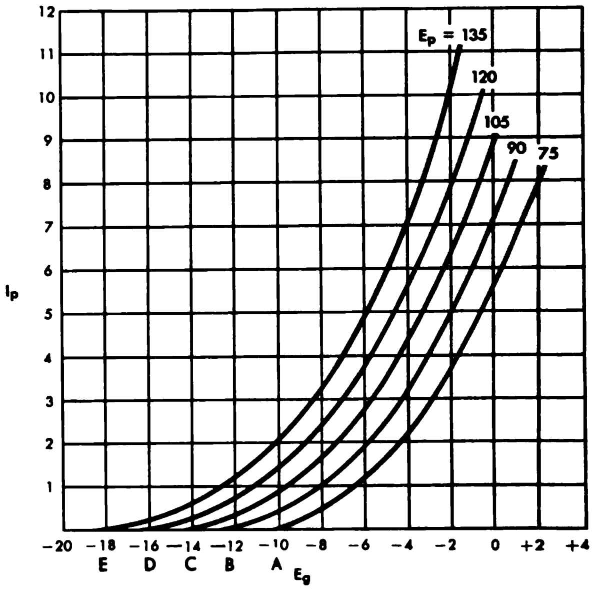
Mutual characteristic curves may be plotted by using a circuit similar to that shown in the illustration below. If E_p is set at 75 volts and E_g is varied in steps of 2 volts from 0 to −12, I_p will vary according to the table of values shown.



Circuit for Plotting Mutual Characteristic Curves

E_g	0	−2	−4	−6	−8	−10
I_p	5.8	3.8	2.2	1.0	0.4	0

In the illustration, A Family of E_g-I_p Curves, curve A results if these values are plotted. Note that plate current is zero when the grid voltage is −10; therefore, cutoff bias with a plate voltage of 75 volts is −10 volts for this triode.



A Family of E_g-I_p Curves

If the process is repeated with E_p set at 90 volts, curve B results. Cutoff for this value of E_p is −12 volts. Repeating the process with plate voltages of 105 volts, 120 volts, and 135 volts results in curves C, D, and E respectively. These curves are known as a family of E_g-I_p curves for the triode being used. Note that as E_p becomes greater, E_g must be made more negative to reach cutoff bias.

Plate Characteristic Curves

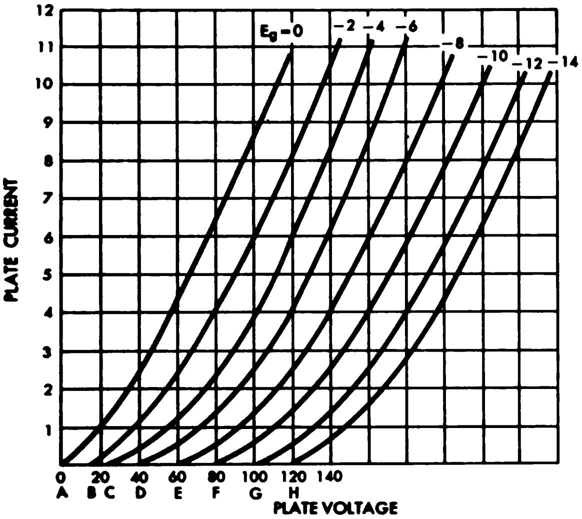
In addition to the E_g-I_p curves, it is also possible to develop a family of E_p-I_p curves, using the circuit shown to the left. If E_g is set at 0 volts and E_p is varied from 0 to 120 volts in steps of 20 volts, the values of I_p shown in the table result.

E_p	0	20	40	60	80	100	120
I_p	0	1.0	2.5	4.4	6.5	8.6	10.9

Plotting these values yields curve A of the family of E_p-I_p curves shown in the illustration below.

If E_g is set at -2 volts and the process repeated, the values obtained yield curve B of the family of E_p-I_p curves. Making E_g two volts more negative each time and repeating the process results in the rest of the curves shown.

Note that these curves, like the family of E_g-I_p curves are almost equally spaced and of the same shape. Both families of curves are known as *static* characteristic curves because they are obtained by applying fixed voltage values directly to the electrodes. Such curves for any of the tubes manufactured may be found in tube manuals. You will use them in the solution of various problems and circuits.



A Family of E_p-I_p Curves

TUBE CHARACTERISTICS

The characteristic curves discussed in the previous section are static and would be of no use in predicting the performance of a tube under operating conditions, were it not for the fact that each tube has certain dynamic characteristics. These characteristics may be used as a method of evaluating the ability of a tube to perform specific functions under operating conditions. Manufacturers of tubes list three such characteristics in their tube manuals: amplification factor (μ), mutual conductance (g_m), and AC plate resistance (r_p).

The Amplification Factor

The amplification factor (μ) is a measure of the relative effectiveness of the grid to that of the plate in controlling plate current flow. It has already been defined as the ratio of the change in plate voltage to the change in grid voltage, required for a given plate current change. This definition may be expressed mathematically as

$$\mu = \frac{\Delta E_p}{\Delta E_g} \text{ (} I_p \text{ constant)}$$

where ΔE_p is the change in plate voltage and ΔE_g is the change in grid voltage. The amplification factor for the 6C5 may be found to be 20 in your tube manual.

The Mutual Conductance

Mutual Conductance, or transconductance, is the ratio of a small change in plate current to the small change in grid voltage which produced it — with the plate voltage kept constant. Expressed in formula form,

$$g_m = \frac{\Delta I_p}{\Delta E_g} \text{ (} E_p \text{ constant)}$$

where g_m is the mutual conductance, ΔI_p is the change in plate current, and ΔE_g is the change in grid voltage. When ΔI_p is in amperes and ΔE_g in volts, g_m is expressed in mhos. Mho is derived from the word ohm spelled backward.

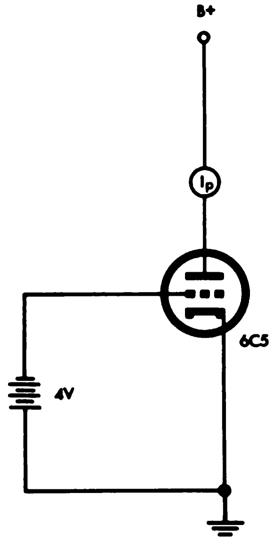
A mho is a relatively large unit for expressing the mutual transconductance of a tube. Manufacturers therefore rate their tubes in micromhos. A mho is equal to one million micromhos. The value of g_m for the 6C5 tube is listed as 2000 micromhos in tube manuals.

The AC Plate Resistance

The AC plate resistance (r_p) is the opposition that a tube offers to AC current flow. It is the ratio of a small change in plate voltage to the change in plate current it produces, with the grid voltage kept constant. Expressed in formula,

$$r_p = \frac{\Delta E_p}{\Delta I_p} \text{ (} E_g \text{ constant)}$$

where r_p is the AC plate resistance expressed in ohms.



Finding the DC Resistance of a Tube

The ratio expressed by the plate resistance is the reciprocal of the slope of the E_p-I_p curve at a particular point. This value depends to some extent on the conditions under which the tube is operating. It is fairly constant, however, along the straight portion of the curve. The tube manual gives the value of r_p for the 6C5 tube as 10,000 ohms for the recommended operating conditions.

Relationship of AC Plate Resistance to Load

You will recall from AC circuit analysis that for maximum power transfer, the impedance of the power source should equal the impedance of the load. For example, in a common radio the vacuum tube in the final stage of amplification may be considered as a source of AC power; the speaker may be considered as the load to which this power is supplied. The AC plate resistance (impedance) of this vacuum tube is usually very large compared to the impedance of the speaker, and little power will be transferred to the load unless an impedance matching transformer is used. It then follows that you must know the AC impedance of the tube as well as that of the speaker so that a transformer with the proper turns ratio may be chosen for maximum power transfer.

Relationship Between μ, g_m, and R_p

The three tube characteristics — μ, g_m, and r_p — are related. Since μ is defined as a ratio of two voltages, it is expressed as a numerical ratio without any unit. However, g_m, which is a ratio of a current to a voltage, is expressed in *mhos;* while r_p, which is a ratio of a voltage to a current, is expressed in *ohms*. An ohm and a mho are reciprocals; therefore, multiplying an ohm by a mho yields unity (or 1).

If g_m (in mhos) is multiplied by r_p (in ohms), the result is a numerical ratio like μ.

$$g_m \times r_p = \frac{\Delta I_p}{\Delta E_g} \times \frac{\Delta E_p}{\Delta I_p}$$

After canceling ΔI_p,

$$g_m \times r_p = \frac{\Delta E_p}{\Delta E_g}$$

But

$$\mu = \frac{\Delta E_p}{\Delta E_g}$$

Therefore, $\mu = g_m r_p$

This formula may be checked by using the values for the 6C5 tubes. From the tube manual, you found that $\mu = 20$, $g_m = 2000$ micromhos, and $r_p = 10{,}000$ ohms.

$$\mu = g_m r_p$$
$$\mu = .002 \times 10000$$
$$\mu = 20$$

If any two of the characteristics are known, the other may be found from this relationship and from the two derived from it: $g_m = \frac{\mu}{r_p}$ and $r_p = \frac{\mu}{g_m}$.

DC RESISTANCE

A tube, like most other circuit components, has DC resistance. You may determine the DC resistance of the 6C5 tube in the circuit shown to the left by measuring the current I_p and then using Ohm's law $R_p = \frac{E_p}{I_p}$.

Instead of measuring the current I_p, however, you may use the $E_p - I_p$ family of curves for the 6C5 shown at the top of the next page. A plate voltage of 150 volts and grid voltage of -4 volts determines point A. The plate current at the point is 6 ma and the DC resistance is found as follows:

$$R_p = \frac{E_p}{I_p}$$

$$R_p = \frac{150}{.006} = 25{,}000 \text{ ohms}$$

Effect of Bias

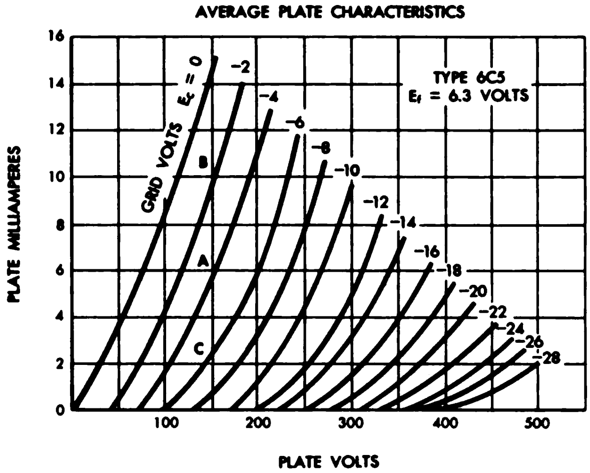
Bias is the negative DC voltage applied to the grid of a tube and is needed to prevent the distortion that results from a positive grid voltage. The addition of bias to the incoming signal will result in a grid voltage which varies below the zero reference line. (See the figure below.)

Another need for bias is to control the DC current flow through the tube. The lower this current level, the less power will be consumed by the amplifier circuit.

The DC resistance of the tube is 25,000 ohms under the conditions shown in the circuit when $E_p = 150$ volts and $E_c = -4$ volts. Suppose you change the grid bias to -2 volts and note the effect this has on the DC resistance. Point B on the family of curves is where the plate voltage is 150 volts and the grid voltage is -2 volts. At this point, the plate current is 10 ma and the DC resistance is

$$R = \frac{150}{.01} = 15{,}000 \text{ ohms}$$

Note that the DC resistance has decreased with a decrease in bias and the plate voltage is unchanged.



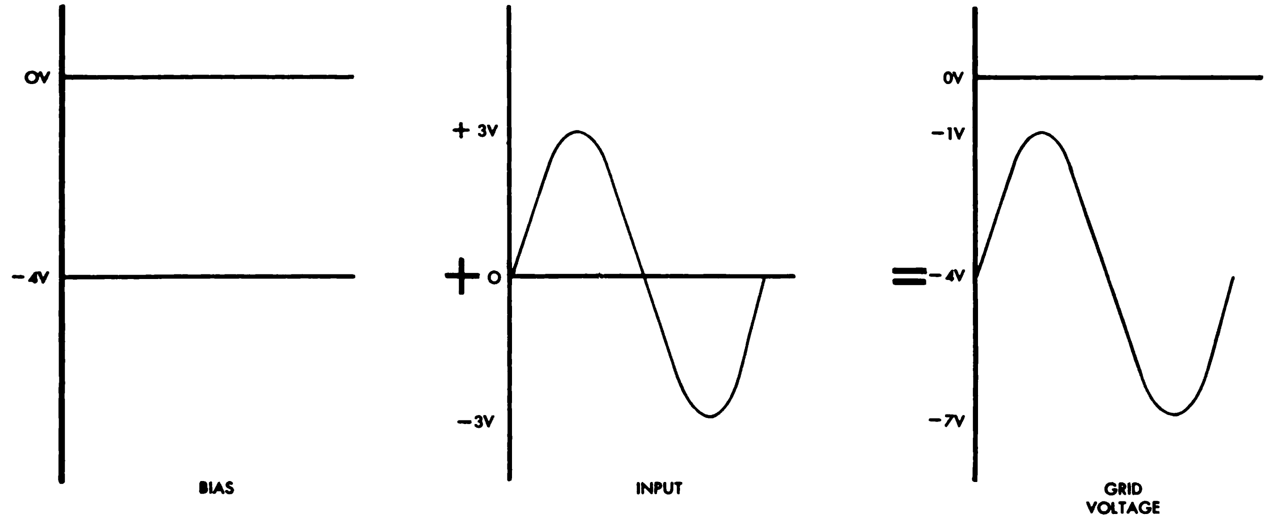
E_p-I_p Curves (6C5)

Suppose you change the bias to -6 volts and keep the plate voltage at 150 volts. These conditions exist at point C of the family of curves. The plate current at this point is approximately 2.6 ma and the DC resistance is

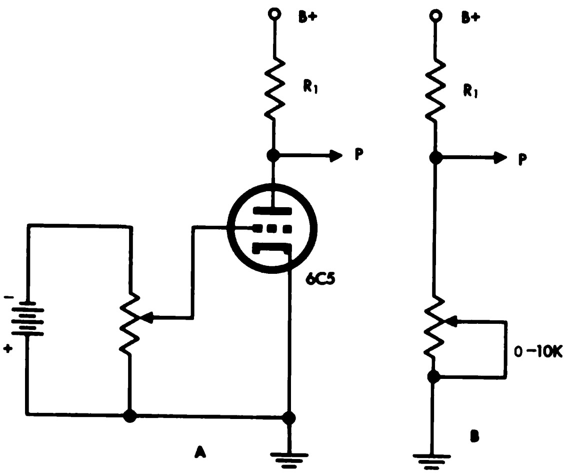
$$R = \frac{150}{.0026} = 57{,}500 \text{ ohms}$$

Note that the DC resistance has increased with an increase in bias with the plate voltage remaining unchanged.

The effect of bias on the plate current and the DC resistance of a tube may be summarized as follows: An increase in bias decreases plate current flow and increases the DC resistance; a decrease in bias increases the plate current



Effect of Bias



The Triode as a Variable Resistor

flow and decreases the DC resistance. These results may be checked by using other points on the family of curves along any constant plate-voltage line.

The Triode—A Variable Resistor

Because its DC resistance varies as its grid voltage is varied, a triode may be used as a variable resistor in many circuits. Compare the two circuits shown above and note the similarity. The tube with the variable bias in diagram A may be considered to be similar to the rheostat in diagram B. As the DC resistance of the tube is increased by making the grid voltage more negative, it has the same effect as increasing the resistance of the rheostat in diagram B. The voltage at P with respect to ground becomes greater in both cases. In similar fashion, making the grid less negative has the same effect as decreasing the resistance of the rheostat and making the voltage less at P. Since the voltage at P is the plate voltage of the tube in diagram A, the following conclusion may be drawn: the plate voltage of a triode decreases as the grid voltage is made more positive and increases as the grid voltage is made more negative. Another way of saying this is — plate voltage and grid voltage are *180° out of phase.*

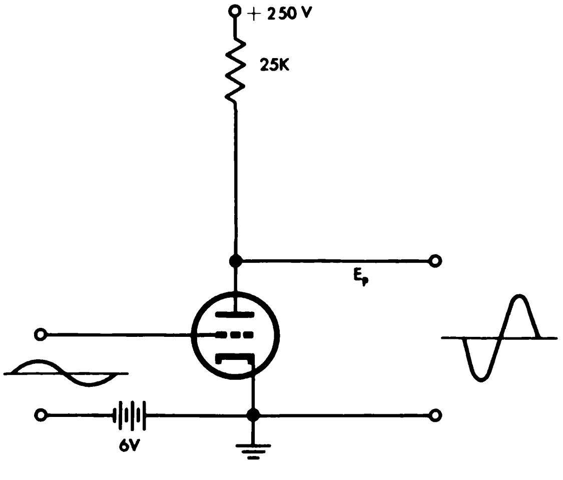
That the plate voltage and the grid voltage are 180° out of phase is also evident if you consider the current flowing through the tube in diagram A. If the grid is made more negative, less current flows through the tube and the voltage drop across R becomes less. The voltage at P therefore becomes greater. With the voltage on the grid made more positive, the reverse situation occurs and the voltage at P becomes less.

AMPLIFICATION

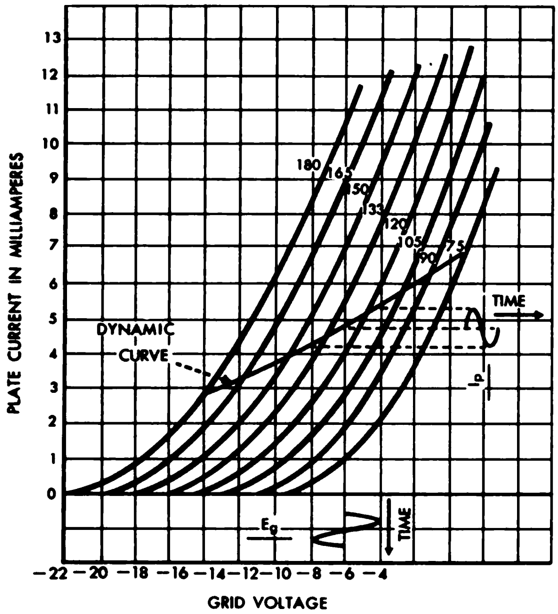
A sine-wave signal applied to the grid of a triode alternately raises and lowers the voltage on the grid. This, in turn, lowers and raises the plate voltage output as shown in the illustration below, Amplifying a Signal. The output voltage is 180° out of phase with the input signal, and of much greater amplitude. Here you see illustrated the fact that a vacuum tube which has a grid has the ability to amplify. The amplitude of the output voltage may be determined by using the E_g-I_p or E_p-I_p characteristic curves. These methods are explained in the following paragraphs.

The Dynamic Curve

The static E_g-I_p characteristic curve shows the effect of a change in grid voltage on the plate current flow with the plate voltage kept constant. When a signal is applied to the grid, the plate voltage in the circuit below varies;



Amplifying a Signal



Plotting a Dynamic Curve

therefore, it is not possible to determine the plate current for each value of grid voltage directly from the static E_g-I_p curve. You can, however, use the static E_g-I_p curves to plot a *dynamic* characteristic curve for any given value of applied voltage and plate-load resistor (R_L). From this dynamic curve the value of plate current for any given grid voltage can be found.

The illustration above shows the dynamic curve for the circuit plotted on the family of E_g-I_p curves for the tube. This dynamic curve is for an applied voltage of 250 volts and $R_L = 25K$. The method of plotting the curve is as follows.

For the plate voltage to be 180 volts, a voltage drop of 70 volts must take place across R_L. To find the current, use Ohm's law.

$$I_p = \frac{70}{25{,}000}$$

$$I_p = 2.8 \text{ ma.}$$

This point is located on the 180-volt curve. The same procedure is followed for all points on the dynamic curve. The results for each point are shown in the table.

E_p (volts)	I_p (ma)	Voltage Drop Across R_L
180	2.8	70
165	3.4	85
150	4.0	100
135	4.6	115
120	5.2	130
105	5.8	145
90	6.4	160
75	7.0	175

If a 4-volt peak-to-peak AC voltage is applied to the grid, the grid voltage varies from −4 volts to −8 volts. From the dynamic curve, you find that the plate current varies from 5.4 ma to 4.3 ma. The plate voltage, therefore, varies from 142.5 volts to 115 volts. These values are found as follows:

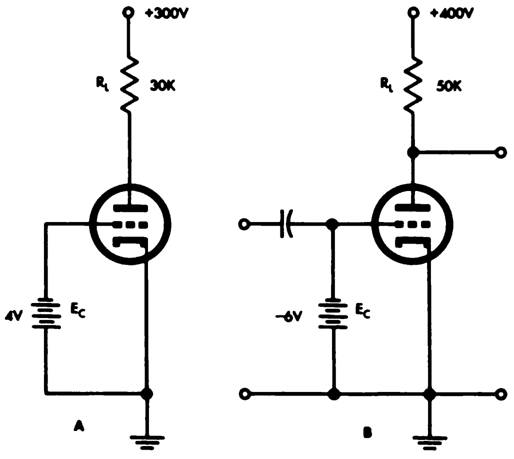
$$E_p = (B+) - IR_L$$

$$E_p = 250 - (4.3 \times 10^{-3} \times 25000) = 142.5 \text{ volts}$$

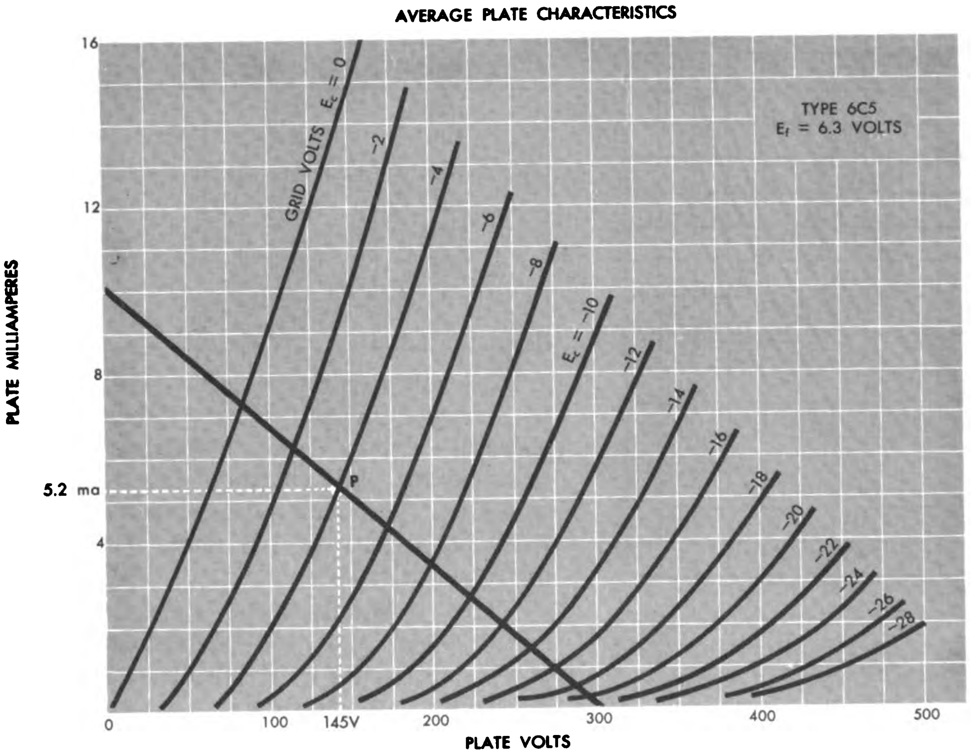
$$E_p = 250 - (5.4 \times 10^{-3} \times 25000) = 115 \text{ volts}$$

The plate voltage therefore varies 27.5 volts.

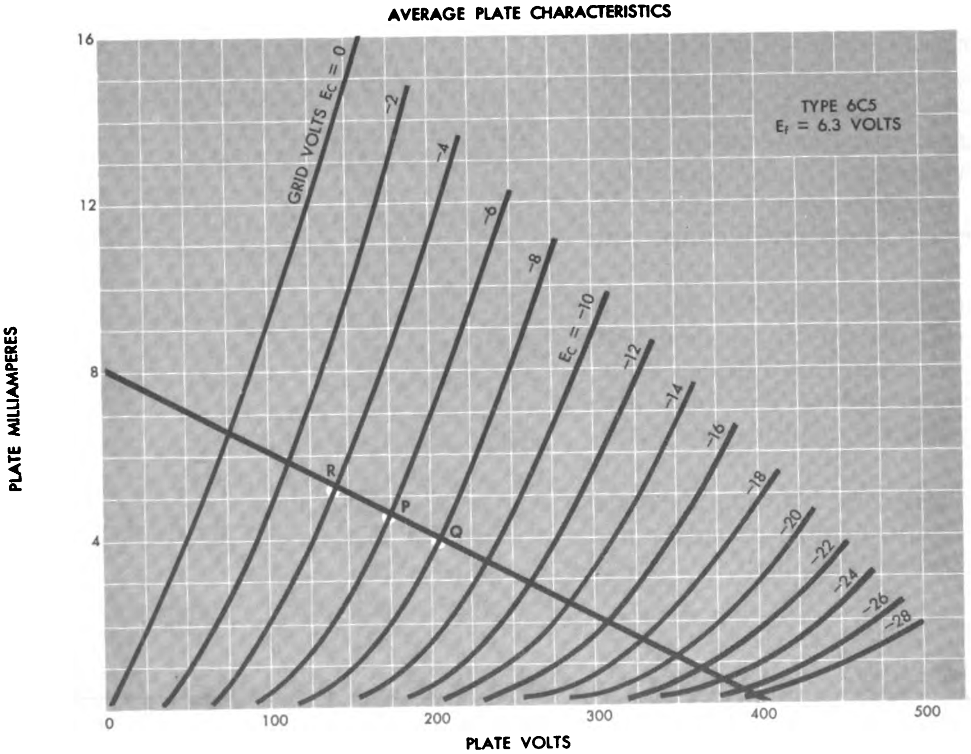
The voltage gain of a stage is defined as the ratio of the output signal to the input signal. The voltage gain of the amplifier for which the dynamic curve was drawn is $\frac{27.5}{4} = 6.9$. However, the amplification factor (μ) of this tube is 10. The actual gain of a circuit is always less than the amplification factor if the tube is used as a voltage amplifier with a resistive load.



Voltage Amplifiers



DC Load Line and Operating Point



Using Load Line for Finding Gain

Load Lines

Many vacuum tube problems may be solved graphically with the aid of E_p-I_p curves and a *load-line* which may be drawn for a given load and a given value of B+. For a purely resistive load, a static or DC load line may be used to determine the output voltage if the tube, load resistor, bias, B+, and input voltages are known.

Using Kirchhoff's law in the vacuum tube circuit shown on page 167 you may set up the following equation:

$$E_p + I_pR_L = B+$$

If $R_L = 30K$ and $B+ = 300$ volts, the equation becomes

$$E_p + I_p\,(30K) = 300$$

This equation has two unknowns (E_p and I_p), and may therefore be plotted as a straight line. Since two points determine a straight line, use the point where $E_p = 0$ and the point $I_p = 0$.

$$\text{When } E_p = 0,\ I_p = \frac{300}{30K} = 10 \text{ ma,}$$

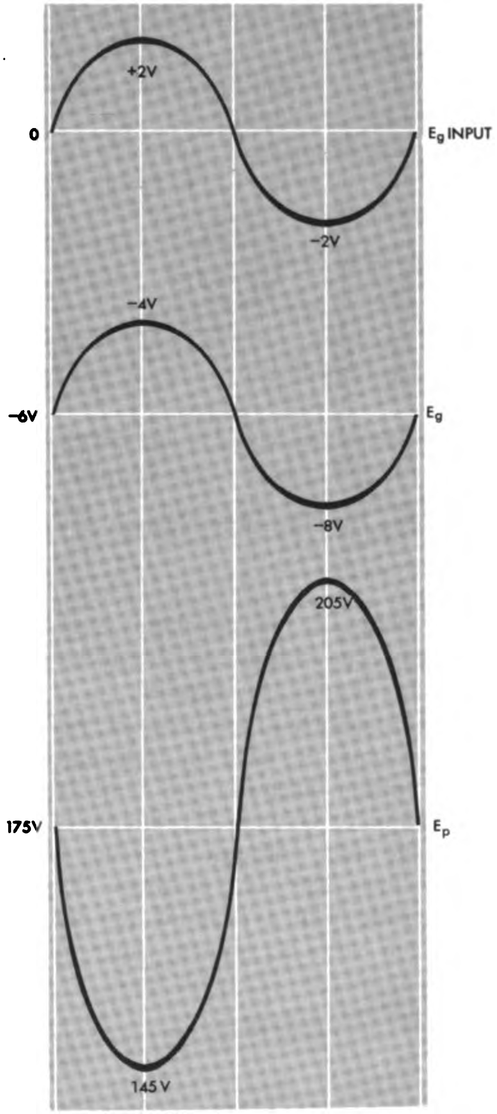
$$\text{When } I_p = 0,\ E_p = 300 \text{ volts}$$

Plot these points on an $E_p - I_p$ family of curves for a 6C5 and join them with a straight line. This line — the solid line shown in the illustration on page 168 — is a DC load line for a 30K load resistor and a B+ of 300 volts.

The *operating point* of a tube determines the values of plate voltage, grid bias, and plate current under static conditions — before a signal is applied. The DC load line is made up of an infinite number of points showing plate voltage and plate current. The operating point of a tube is somewhere on this load line. To find this point requires the intersection of the load line with another line which gives the bias voltage — the bias line.

For this circuit the −4 volt bias line is used. The intersection (point P) of the −4 volt bias line with the load line is the operating point. Here, $E_p = 145$ volts, $I_p = 5.2$ ma, and $E_c = -4$ volts.

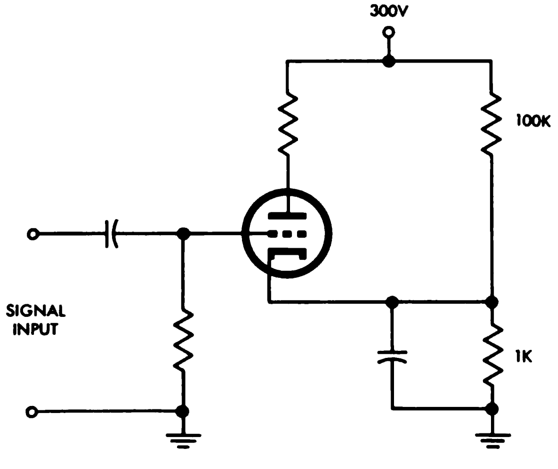
In the circuit shown at B, $R_L = 50K$, $B+ = 400$ volts, and $E_c = -6$ volts. The load line for this circuit is drawn on the lower family of curves for the 6C5 tube shown on page 168. Since the tube is operating with a bias of −6 volts, the operating point is somewhere on the −6-volt bias line. The intersection (point P) of the load line with the bias line is the operating point for the conditions given. Here $E_p = 175$ volts, $I_p = 4.6$ ma., and $E_c = -6$ volts.



Input and Output of Voltage Amplifier

If E_c were increased to −8 volts, the operating point would become point Q at which $E_p = 205$ volts and $I_p = 4$ ma. A decrease of the bias voltage to −4 volts results in an operating point at R with $E_p = 145$ volts and $I_p = 5.2$ ma. Note that the operating point remains on the load line for each case.

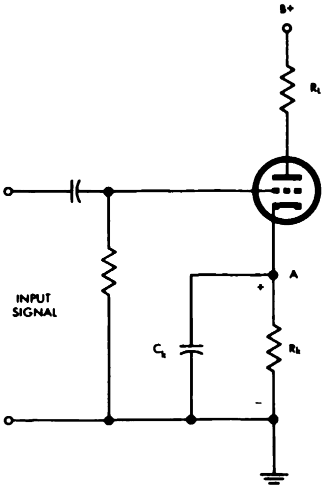
Apply a 4-volt peak-to-peak signal to the circuit as shown to the left. The load line and



Obtaining Fixed Bias from Voltage Divider

operating point (P) for this circuit have already been drawn. The voltage on the grid therefore varies from −4 volts to −8 volts.

When the tube has −4 volts on the grid, it operates at point R, with $E_p = 145$ volts; when the grid voltage becomes −8 volts, the tube operates at point Q, with $E_p = 205$ volts. As the grid voltage goes through a cycle, the plate voltage follows 180° out of phase. The plate voltage varies from a peak of 145 volts to a peak of 205 volts as shown. Since the plate voltage varies 60 volts while the grid voltage signal is only 4 volts, the gain of the stage is 15. Note that μ for the 6C5 is 20 and that the gain of the stage is less.



Cathode Bias

OBTAINING BIAS

Bias was defined as a DC voltage which makes the grid of a tube negative with respect to the cathode. If the biasing voltage is applied directly to the grid or to the cathode from a battery or other external source, the bias is called *fixed bias*. If the biasing voltage is mainly a result of current flowing in the tube, the bias is called *self-bias*.

Fixed Bias

The type of fixed bias in which a battery voltage is applied directly to the cathode or grid has been discussed. The circuits on page 161 are examples of bias obtained from a C battery. Fixed bias is often obtained in radio or radar circuits without the use of a C battery, however. As shown in the circuit to the left above, a positive voltage from a voltage divider is applied to the cathode. The bias in this case is determined by the current in the voltage divider and the current flowing through the tube under static conditions. These currents flowing through the resistor R_1 cause a voltage drop which makes the cathode positive with respect to the grid.

Cathode Bias

A common method of developing bias from the current flowing through the tube is shown in the circuit to the left. This type of self-bias is called cathode bias because the bias that is developed results from the current flow through the cathode resistor. As the tube conducts, a voltage (I_pR_k) is developed across the cathode resistor. For example, if 2 ma. flows through a 2K resistor, 4 volts is developed. The current flows from ground through R_k and the tube; therefore, A is positive with respect to ground. Since the grid is at ground potential, the grid is negative with respect to the cathode, thus biasing the tube.

Under static conditions the tube stabilizes itself with some bias on the grid. However, if a signal were applied, the current flow, and therefore the bias, would increase and decrease with the signal. This would cause degeneration — a reduction in the amplitude of the voltage swing at the plate of the tube. A large capacitor (C_k) is placed across the resistor (R_k) to prevent degeneration by stabilizing the bias when an AC signal is applied. The AC variations meet very low impedance through C_k while the DC current flows through R_k; this keeps the bias almost constant.

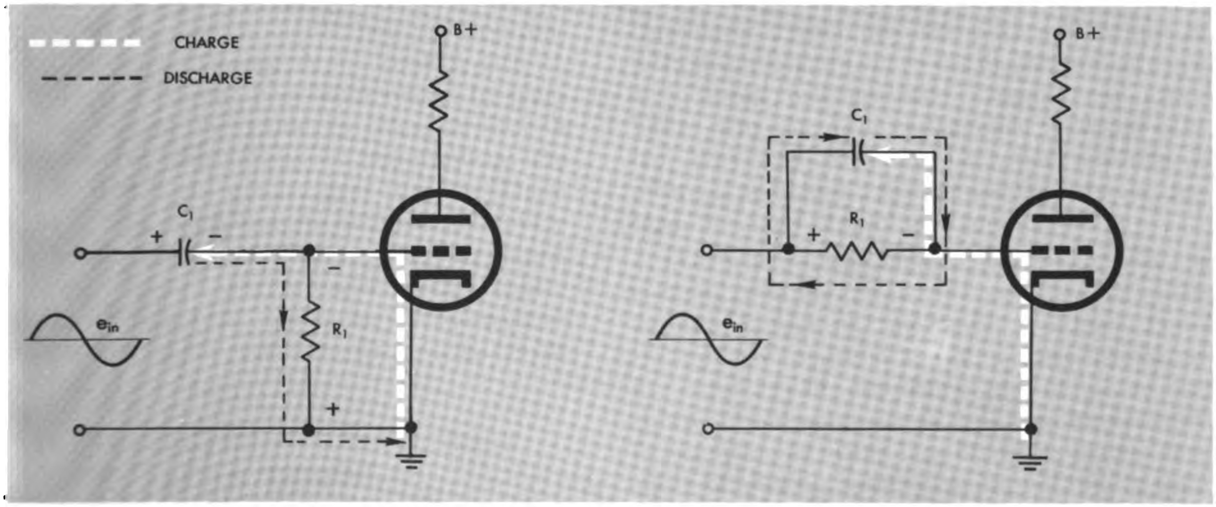
Note that a tube *cannot be cut off* with cathode bias as the only type of bias. This is due to the fact that bias is developed only when the tube conducts; if there is no conduction, no bias is developed. Therefore, the tube must conduct and the amount it conducts depends upon the bias developed.

Grid-Leak Bias

Another form of self-bias is called *grid-leak* bias. This type of bias depends upon the input signal. Two circuits using grid-leak bias are shown in the illustration.

The action is similar in both cases and is explained as follows: When an AC signal is applied the grid draws current on the positive alternation. The condenser C_1 charges, as shown by the white-dotted arrow. When the signal goes through the negative alternation, the condenser C_1 tries to discharge. The only path through which the electrons can flow is through the resistor R_1, as shown by the black-dotted arrow. A negative voltage is developed across R_1 which biases the tube. Since the resistor R_1 is very large, the condenser discharges slightly and only a very small current flows or leaks through. The resistor is therefore called a grid-leak resistor.

The main disadvantage of grid-leak bias is that bias is developed only when a signal is applied to the grid. If the signal is removed for any reason, the tube conducts very heavily and may be damaged. This type of bias is used frequently in oscillator circuits. *Combination bias* is the term applied to the use of grid-leak bias in conjunction with fixed or cathode bias. This combination provides the advantages of grid-leak bias with an added safety precaution in case the signal is removed.



Grid Leak Bias

USES OF TRIODES

The use of a triode as an amplifier has been briefly discussed. There are a great many different types of amplifiers, each of which will be studied in greater detail later. The tubes used in all of these amplifiers have control grids; in many applications, they are triodes.

In addition to its use as an amplifier, the triode has countless other applications in electronics. You will become acquainted with a few of these special applications in this course. Some applications of triodes which you will study are: limiters, clippers, detectors, oscillators, sweep generators, sawtooth generators, voltage regulators, control tubes, etc. Although these applications of the triode may be only meaningless words to you at this time, they are listed to give you an idea of a few of the many uses made of triodes. The insertion of the control grid in the vacuum tube by Lee DeForest opened vast possibilities in the field of electronics that have not as yet been fully realized

Chapter 21

Tetrodes and Pentodes

The ability of a triode to amplify depends on the effectiveness of the control grid as compared to that of the plate in controlling plate current flow. This effectiveness of the control grid to control plate current may be increased by placing the control grid closer to the cathode and by increasing the density of the mesh of the wires of which the control grid is made. These factors set physical limitations upon the ability of a triode to amplify. Another factor which limits the use of triodes under certain conditions is *interelectrode capacitance* — the capacitance which exists between the electrodes of the tube.

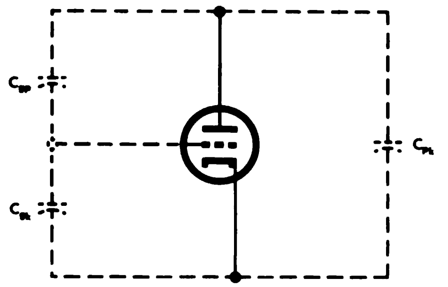
INTERELECTRODE CAPACITANCE IN A TRIODE

In your study of AC you learned that two conductors separated by a dielectric form a capacitor. You also learned that the capacitance of a capacitor depends upon the size of the conductors, the distance between them, and the dielectric.

A triode has three electrodes (conductors) separated by a dielectric (vacuum). The conditions for capacitance therefore are present in a triode and, as shown in the illustration on page 174, three interelectrode capacitances exist. These are: C_{gk}, the capacitance between control grid and cathode; C_{pk}, the capacitance between plate and cathode; and C_{gp}, the capacitance between control grid and plate.

The grid-to-plate capacitance (C_{gp}) offers a parallel path for plate current. The AC plate current divides at the plate; one part flows through the load and the other part returns to the input circuit. Returning part of the output signal to the input circuit is known as *feedback*. In receivers, feedback may have the undesirable effect of causing howling, squealing, and a general lowering of sensitivity.

Since the interelectrode capacitance C_{gp} is always present, feedback is also always present. The effects of feedback, however, depend upon the amount of feedback. If the feedback is small, its effects are not noticeable.



Interelectrode Capacitances of a Triode

The reactance of C_{gp} determines the amount of voltage fed back to the grid. A large reactance causes the feedback to be small. As you recall, capacitative reactance is expressed in formula form as $X_c = \frac{1}{2\pi fC}$. The value of the interelectrode capacitances is small (being measured in micromicrofarads) because of the small area of the electrodes. At audio frequencies, f in the formula is also small; therefore, the reactance is large enough to make the amount of feedback negligible. At radio frequencies however, f is large enough to lower the reactance and cause an excessive amount of feedback to occur. The higher the frequency, the greater is the amount of feedback which occurs.

TETRODES

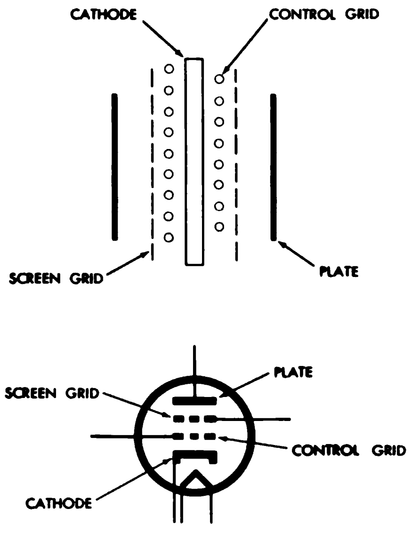
In the late 1920's, a new type vacuum tube was introduced which successfully reduced the plate to control grid capacitance to such an extent that the amount of feedback was negligible for most radio-frequency uses. This type of tube is called a tetrode because it has four electrodes — the three of the triode plus a screen grid located between the control grid and the plate. The screen grid or screen is made in gridiron form and is similar in construction to the control grid. This type of construction permits the flow of electrons to reach the plate. The purpose of the screen is to shield the plate from the control grid electrostatically. The schematic representation and construction of the tetrode is shown in the illustration to the right.

Typical Tetrode Circuit

To learn how a tetrode reduces interelectrode capacitance and minimizes feedback, examine the illustration, Simplified Circuit of a Tetrode Amplifier, page 175. The screen grid is connected to B+ through a resistor R_s, called a *dropping resistor*, and to ground through a *bypass capacitor* C_s. As in the case of the triode, interelectrode capacitances exist between the electrodes. Two of these are shown dotted in the illustration. These are C_{sp}, the capacitance between screen grid and plate, and C_{gs}, the capacitance between control grid and screen grid. The value of the bypass capacitor C_s is chosen to present a very low reactance (practically a short) at the frequency of the signal.

The space current from the cathode flows in two circuits, the plate circuit and the screen grid circuit. Trace the path of the AC circuit of the screen grid first. The current which reaches the screen can flow in either of two paths: through C_{gs} to the input circuit to ground, or through C_s to ground. Since C_s is chosen to be almost a short circuit, most of the current flows through it. Therefore, there is no sizable feedback from the screen grid to the control grid.

Now trace the path of most of the space current which continues to the plate. This AC current also has two paths — either through C_{sp} or through the load. The portion which flows through C_{sp} and reaches the screen also has two paths through C_{gs} or through C_s. Most of the current flows through the low impedance of C_s. Therefore feedback to the input circuit is minimized and its effects avoided.



Construction and Schematic Representation of a Tetrode

Secondary Emission

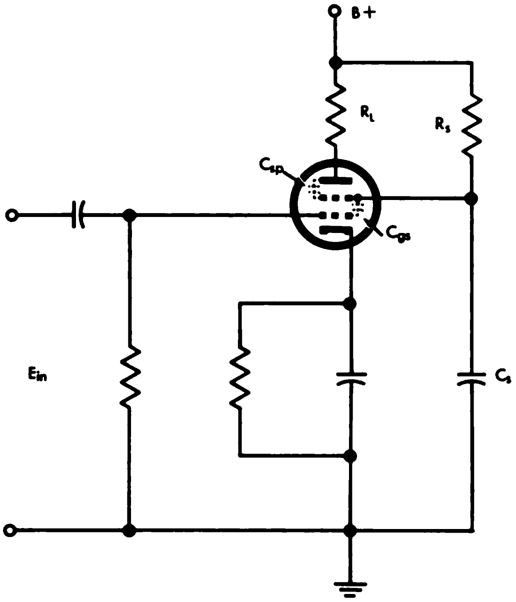
Secondary emission was discussed as one type of emission on page 151. It occurs if electrons are accelerated and allowed to strike a metal plate with sufficient force to release free electrons and emit them into the space surrounding the plate. The number of free electrons emitted depends upon the number and velocity of the moving electrons and the number of free electrons in the substance they strike.

Because the screen grid of a tetrode is operated at a positive potential, it causes the electrons emitted from the cathode to accelerate as they travel toward the plate. These accelerated electrons strike the plate with sufficient velocity to cause secondary emission from the plate into the space between the screen grid and the plate. The electrons emitted in this fashion travel toward the electrode exerting the greater force on them. The positive potentials on the plate and on the screen grid determine which electrode has the greater effect and attracts the greater number of these electrons. At low values of plate voltage, the secondary electrons form a cloud or space charge between screen grid and plate. This space charge repels electrons from the plate toward the positive screen grid. An examination of the family of E_p-I_p curves for a typical tetrode shown below will give you a better understanding of what effect secondary emission has.

Characteristic Curves

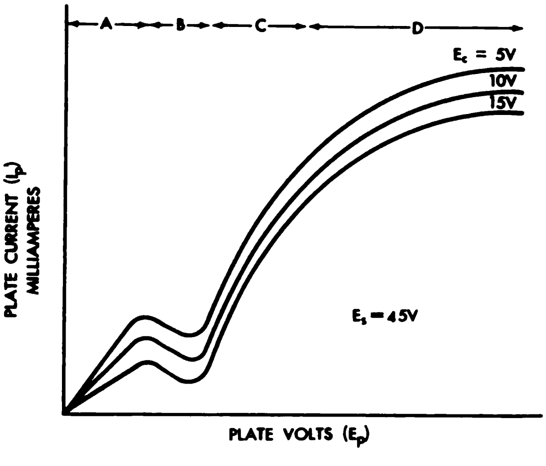
Each curve of the family of curves is obtained by keeping the control grid voltage and screen grid voltage constant and plotting the variation in plate current as the plate voltage is varied. A different value of grid voltage is used for each curve of the family of curves; however, the screen grid voltage is kept at the same value for the entire family of curves. Since the screen grid is positive with respect to the cathode, electrons are attracted to it and a small current flows in the screen grid circuit. The screen current is small because of its open-mesh construction and the force of attraction on the emitted electrons exerted by the plate voltage. Each curve has four regions as follows:

REGION A. At low values of plate voltage, the plate current increases with an increase in plate voltage. In this region, the screen grid has the same control on plate current flow as the plate has in a triode. As the plate voltage is increased from zero volts, more electrons are attracted toward it.

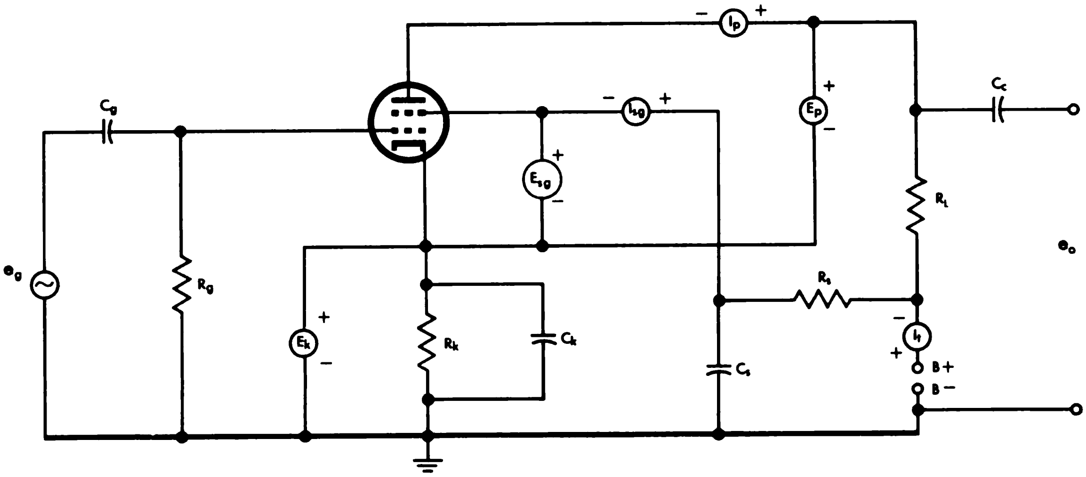


Simplified Circuit of Tetrode Amplifier

REGION B. At slightly higher plate voltages — with the screen grid higher than the plate voltage — the plate current decreases with an increase in plate voltage. Because of this effect, this region is called the *negative resistance* or *dynatron* region of the tetrode. This is the region in which the effect of secondary emission is evident. The secondary electrons are attracted to the screen grid because it is at a higher positive potential than the plate. This increases the screen grid current. The total space current



Family of E_p-I_p Curves Typical of a Tetrode



Tetrode Amplifier

remains approximately the same since the decrease in plate current is accompanied by the increase in screen grid current.

REGION C. From the point on the curve where the plate voltage equals the screen grid voltage, plate current and plate voltage increase together. In this region all the secondary electrons return to the plate instead of entering the screen.

REGION D. For greater values of plate voltage, a large increase in plate voltage results in a *slight* increase in plate current. This resembles plate saturation in the triode. The plate has little control on the flow of electrons from the cathode because the screen grid shields it. The screen grid voltage determines where this region begins.

Because of the effects of secondary emission, tetrodes are operated with high plate voltages — in region D of the curve. When operated in this manner, the values of r_p and μ for the tetrode are high. The value of g_m is about the same as that of a triode operating with the same plate current.

USES OF TETRODES

Tetrodes are seldom used in receivers because pentodes will do the same job for most applications and are less critical in voltage requirements. However, tetrodes are used to a great extent as high power amplifiers in transmitters.

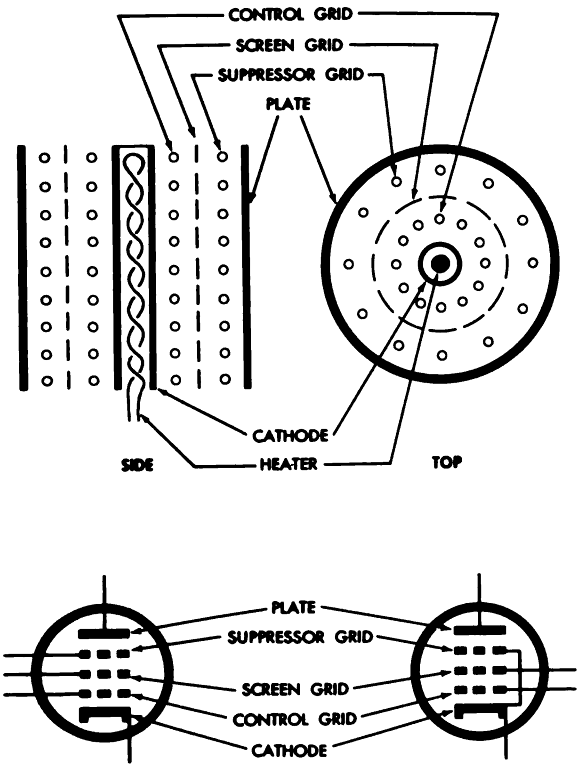
A typical tetrode amplifier circuit is shown in the illustration above. The proper placement of meters for measurement of voltages and currents is also indicated. Note the screen-grid dropping resistor (R_s) in the screen-grid lead. The purpose of the dropping resistor in a tetrode circuit is to make it possible to operate the plate at one potential and the screen grid at a lower potential with only one power supply. The purpose of capacitor C_s is to keep the voltage drop across R_s constant by bypassing the current variations.

PENTODES

Although the addition of the screen grid makes the tetrode a better amplifier, it restricts the use of the tetrode to operation at higher plate voltage. This is due to the fact that the screen grid operating at a positive potential draws excessive current when the plate potential is not high. This excessive screen grid current accompanied by a decrease in plate current is caused by secondary emission from the plate. The pentode — a five element tube — eliminates this fault by adding a third grid, called a *suppressor grid.*

Construction of a Pentode

The construction and schematic diagram of a pentode are shown in the next illustration. Like the other two grids, the suppressor grid is made of spaced wire so that electrons pass



Construction and Schematic Representation of a Pentode

through easily. The suppressor grid is located in the path of the electrons between the screen grid and the plate. It is usually connected to the cathode either internally or externally. The suppressor grid has little effect on the primary electrons emitted from the cathode and accelerated by the positive screen grid. However, the electric field between the plate and the suppressor grid is in such direction as to repel secondary electrons back to the plate. The suppressor grid does not stop secondary emission; it merely prevents the secondary electrons from hitting the screen grid.

Characteristic Curves of a Pentode

The effect of the suppressor grid may be seen from an examination of the family of E_p-I_p curves of a pentode shown. Each curve is obtained in the same manner as for the tetrode — by keeping the bias and screen grid voltages constant and plotting the variations in plate current as the plate voltage is varied. The family of curves is obtained for different values of fixed bias with the screen-grid voltage kept constant throughout.

Comparing Tetrodes and Pentodes

Compare the characteristic curves of tetrodes with those of pentodes and note these differences:

The rise of plate current, at lower values of plate voltage, is more rapid in pentodes.

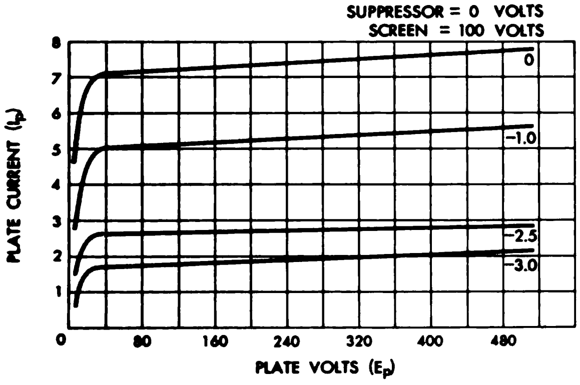
The negative resistance region of the tetrode is eliminated in pentodes.

The straight portion of the curve is greater for pentodes than it is for tetrodes.

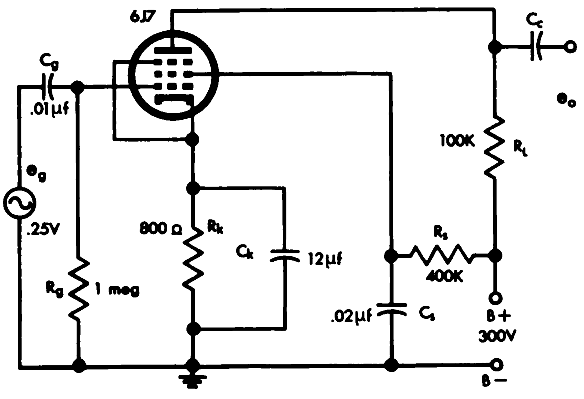
The straight portion of the pentode characteristic curve has less slope than that of the tetrode. This fact means that a greater change in plate voltage is required to obtain a noticeable change in plate current.

USES OF PENTODES

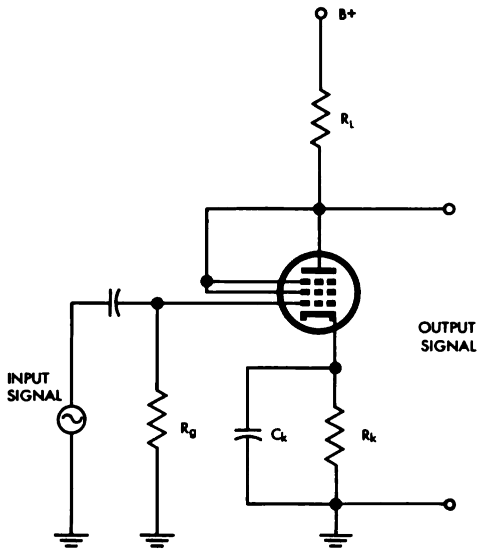
The pentode has almost completely replaced the tetrode as an amplifier in most applications because it has the following advantages:



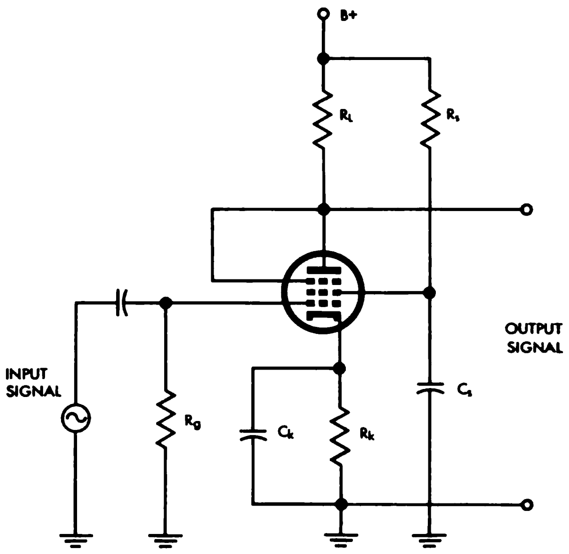
Family of E_p-I_p Curves of a Pentode



Pentode Amplifier



Pentode Connected as a Triode



Pentode Connected as a Tetrode

(1) A greater value of μ and r_p.

(2) Less interelectrode capacitance from plate to control grid because of the presence of an additional shield (the suppressor grid) between them.

A typical pentode amplifier is shown in the illustration on the previous page. Just as in the tetrode circuit, the screen grid is connected to B+ through a dropping resistor (R_s), and a bypass capacitor (C_s) is used. The third grid (suppressor grid) is connected to the cathode externally in this amplifier.

A pentode is often used as a triode by connecting the screen and suppressor grids to the plate as shown. This arrangement changes the tube characteristics so that they compare favorably with those of an ordinary triode. Tube manuals usually list the characteristics of many pentodes for operation as a triode amplifier as well as for pentode operation.

You may also connect a pentode as a tetrode. For this type of connection the suppressor grid is connected to the plate. This use of a pentode is not very common.

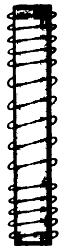
Chapter 22

Special Purpose Tubes

Some tubes have special features which enable them to perform a particular function more efficiently than the tubes discussed so far. Although these tubes may also be classified by the number of electrodes they contain, they are usually identified by the special feature or function they perform. Among these special purpose tubes are remote-cutoff pentodes, beam power tetrodes, multiunit tubes, ultra-high-frequency tubes, and gas tubes.

REMOTE-CUTOFF TUBES

The control grids of most tubes are constructed so that the grid wire, the grid spacing, and the distance between the cathode and the grid are uniform throughout the length of the control grid. As a result, a bias that is sufficient to stop current flow from one point of the cathode cuts off current from all points on the cathode. The characteristic E_g-I_p curves of the tubes studied approach the grid-voltage axis sharply, and these tubes are said to have *sharp-cutoff* grids. The grid wire in a remote-cutoff tube is wound so that the spacing between the turns is not uniform. As shown in the illustration on page 180, the spaces at the two ends of the grid structure are smaller than those in the center. Consequently, a bias which is of sufficient magnitude to cut off the flow of electrons from one part of the cathode may not be of sufficient magnitude to affect other parts of the cathode in the same way. Therefore, the cutoff bias is reached at a more gradual rate, and comparatively large values of bias are required to cut off plate current completely. Compare the E_g-I_p curve of the remote-cutoff tube with that of the ordinary pentode in the illustration on page 180.



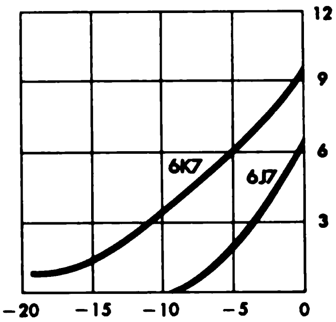
Control Grid Winding of Remote-Cutoff Tube

The amplification obtained, when remote-cutoff tubes are used, varies with the value of bias. An increase in bias causes a decrease in amplification; a decrease in bias causes an increase in amplification. This characteristic of remote-cutoff tubes is used to advantage in avc (automatic volume control) circuits. The bias is made to vary with the strength of the signal applied. This causes the stronger signal to be amplified less than the weaker signal, and an output of more constant level results. Remote-cutoff tubes are therefore known also as *variable-mu* and *supercontrol* tubes.

BEAM POWER TUBES

Many vacuum tubes must be capable of handling large values of current; such tubes are commonly known as power amplifiers, as distinguished from voltage amplifiers. The plate of a voltage amplifier, such as the 6K7, draws a maximum of about 10 milliamperes of current, while a power amplifier, such as the 6L6, may draw as much as 250 milliamperes of plate current. The construction of power amplifiers therefore differs greatly from that of voltage amplifiers in the following respects: the electrodes are larger and of more rugged construction, provisions are made for the dissipation of heat, the spacing between grid-wire turns is greater (to more readily permit the flow of greater quantities of electrons), and the emitting ability of the cathode is greater.

The beam power tetrode is used in many circuits as a power amplifier. The beam power tetrode has the characteristics of a pentode because of the unusual arrangement of its electrodes, as shown at the top of page 181. The control grid and screen grid are wound in such a way that the wires of one lie in the shadow of the wires of the second. The electrons therefore flow in narrow beams between the wires of the control and screen grids. This reduces the amount of screen current (which is wasted power) and inscreases the amount of plate current. The concentration of electrons in the beam is so great that any secondary electrons emitted by the plate are immediately returned to the plate by the negative charge of the beam. Thus, the beam performs the same function as the suppressor grid does in the pentode. Beam forming plates, which are connected internally to the cathode, concentrate the electron beam even more before it reaches the plate.



E_g-I_p Curves for Sharp and Remote-Cutoff Tubes

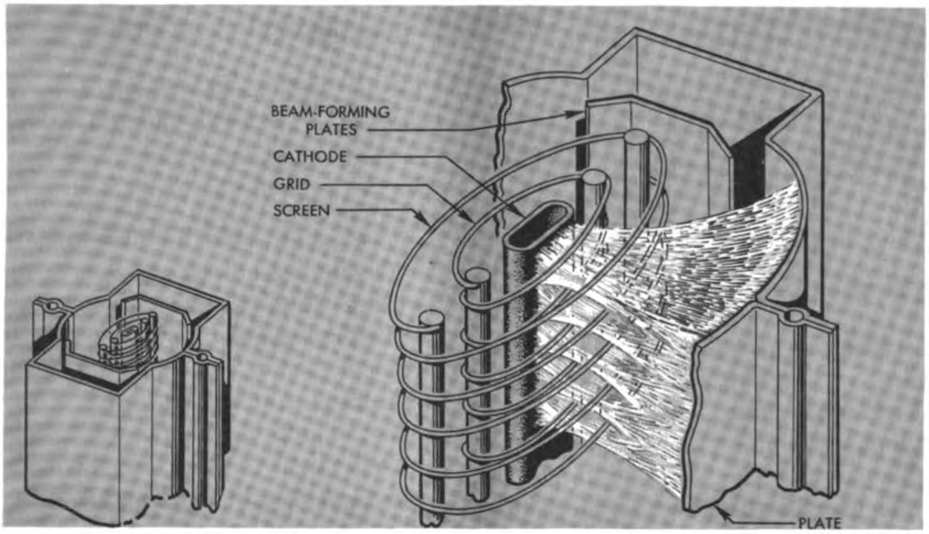
The E_p-I_p curves of beam power tubes are similar to those of pentodes and they are often considered to be a pentode. Schematic representations of beam power tubes are illustrated at bottom of page 181.

MULTIELECTRODE TUBES

Multielectrode tubes contain more than five electrodes. The electrodes added are usually additional grids. A six-electrode tube with four grids is a *hexode;* a seven-electrode tube with five grids is a *heptode;* and an eight-electrode tube with six grids is an *octode.* The five-grid (pentagrid) tube is the most common type of multielectrode tube.

The schematic drawing of a 6A8 pentagrid tube is shown in the illustration on page 181. It has an indirectly heated cathode, a plate, and five grids. The grids are identified by numbers 1 to 5 depending upon their distances from the cathode.

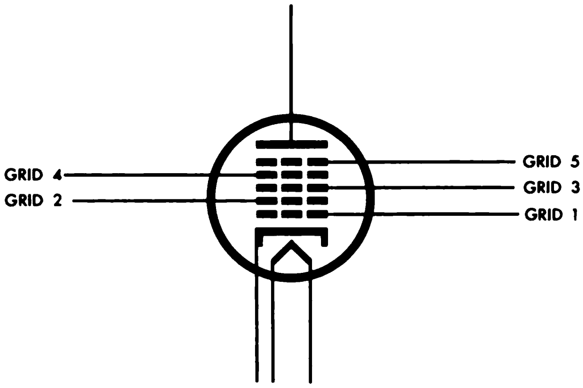
The additional grids make it possible to feed and mix different signals in this tube. If grid 2 is made positive and grid 1 treated as a control grid, the cathode and these two grids may be considered to be a triode. If grids 3 and 5 are connected together and grid 4 treated as a



Construction of Beam-Power Tetrode

control grid, the cathode, plate, and grids 3, 4, and 5 may be considered to be a tetrode. Thus signals may be fed in on grids 1 and 4 and mixed in the tube.

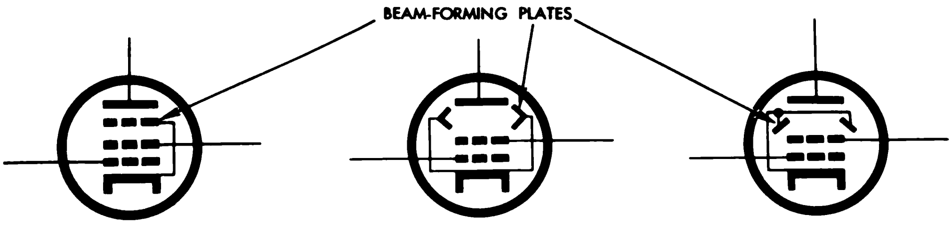
Heptodes (or pentagrid) tubes are used as mixers in superheterodyne receivers. When the triode portion is connected as an oscillator and mixed with an incoming RF signal to the tetrode portion, the tube is known as a *pentagrid converter*. These uses of heptodes will be discussed more fully when you study the first detector of the superheterodyne receiver.



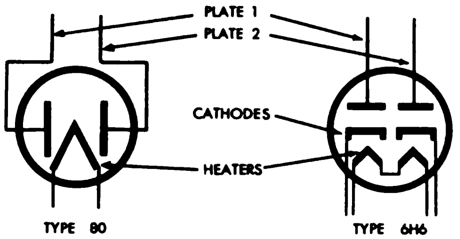
Schematic Representation of 6A8 Tube

MULTIUNIT TUBES

Two or more complete tubes are often built into one envelope to save space and make it possible to use fewer and shorter leads. Another advantage that is gained is that coupling between circuits and unwanted feedback is reduced to a minimum. In some instances, the tubes within the one envelope share the same cathode; however, they differ from the multielectrode tube in one important respect. In the multielectrode tube, there is one stream of electrons emitted from the cathode which is acted upon by all the electrodes. In the multiunit tube, on the other hand, the stream of electrons emitted by the cathode (or cathodes) divides into two or more paths, and each path is through its own unit.



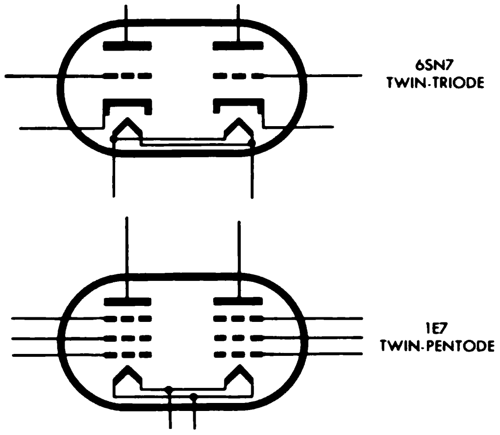
Schematic Representation of Beam-Power Tubes



Schematic of Duo-diodes

Multiunit tubes are constructed so that the electron stream in one unit is not affected by the electrodes of the other units. Although they are built within one envelope, each unit may be treated as a separate tube. Some examples of multiunit tubes are shown above.

The 80 and 6H6 tubes are examples of two diodes in one envelope. Such tubes are known as duo-diodes, twin diodes, or duplex diodes. The 80 tube has a common, directly heated cathode and two plates. This tube is often used as a full-wave rectifier. The envelope of the 6H6 tube contains two complete diodes with indirectly heated cathodes.



Schematic of Twin Triode and Twin Pentode

Some tubes contain two complete triodes in one envelope. A common example of a twin triode is the 6SN7 tube which is shown schematically above. This tube has numerous uses in radio and radar circuits as two separate amplifiers. Twin pentodes like the 1E7 are useful as power amplifiers in push-pull circuits.

Multiunit tubes are not limited to units of the same type only. There are many examples of different types of tubes within one envelope. In the illustration on the next page are shown a diode-triode, a diode-pentode, a duplex-diode triode, a duplex-diode pentode, a triode-pentode, and a diode-triode pentode. Numerous other combinations are possible and are found among the tubes manufactured.

THE TRANSISTOR

One of the newest of important developments in the field of electronics is known as the transistor. The transistor, the name of which was coined from "TRANSfer" and "resISTOR," is a miniature substitute for the triode vacuum tube; it acts as an amplifier. Throughout the development of electronic devices, the tendency has been to add more and more vacuum tubes. The number has been kept within limits only by the use of vacuum tubes which do the work of two or more tubes of simpler construction. The transistor is the first device found which is a true substitute for vacuum tubes in many types of electronic equipment.

The main advantages of the transistor over the vacuum tube are:

Extremely lower power requirements.

Smaller space and weight.

Longer life.

These advantages may be more clearly defined by the following examples:

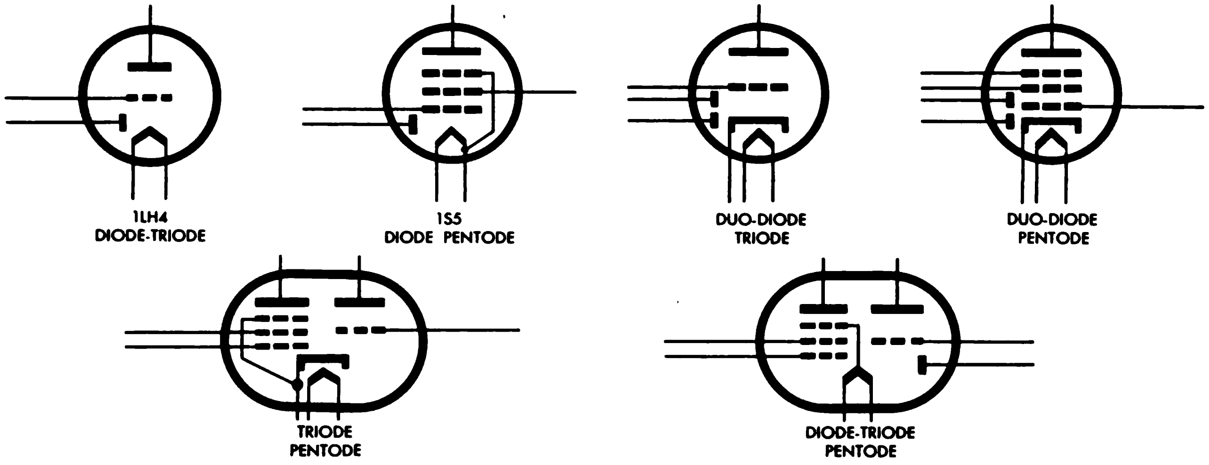
The power saving resulting from the substitution of transistors for vacuum tubes in electronic computers is as high as 1000:1.

The size of present transistors is approximately the same as the size of an ordinary pencil eraser.

Transistors now in use have a life expectancy of over 100,000 hours. In addition, transistors are capable of withstanding physical shocks many thousands of times greater than the force of gravity.

There are many types of transistors now in use or under development. A brief explanation of one of these, the point-contact type, will suffice in a text of this scope.

The point-contact transistor in its present form is a metal cylinder, $\frac{3}{16}$ inch in diameter and $\frac{5}{8}$ inch in length, inside of which is a small piece of germanium and two cat-whisker contacts .002 inch apart. Germanium is one of a number of elements known as semi-conductors,


1LH4
DIODE-TRIODE

1S5
DIODE PENTODE

DUO-DIODE
TRIODE

DUO-DIODE
PENTODE

TRIODE
PENTODE

DIODE-TRIODE
PENTODE

Schematic Representation of Other Multiunit Tubes

which under normal conditions conduct very little electric current. When a small positive potential is applied through a cat-whisker contact on the surface of the germanium, an area around the contact becomes capable of greatly increased conductivity. Small changes of current in the circuit from the first cat-whisker to the base on which the germanium is fastened cause larger changes in the current in the circuit from the second cat-whisker to the base. Voltages are applied to the first and second cat-whiskers in a manner similar to the application of voltages to the grid and plate of a triode except that the polarities are reversed.

Transistors under development have a voltage amplification factor as high as 1000:1 and an upper frequency limit of 600 MC. Noise levels, which in the past have been excessive, are now comparable to the best vacuum tubes.

The possibilities foreseen for the use of transistors in aircraft are truly tremendous. In airborne electronics applications, size, weight, and power consumption are overwhelming considerations. A modern bomber requires about 30 kilowatts to power its electronic equipment. It is difficult to predict how much the transistor will do to reduce this load, with consequent space and weight saving not only in the electronic equipment itself, but in the present power-generating apparatus. The complete transistorization of many electronic black boxes — communications receivers, navigational instruments, digital computers, control and guidance apparatus, radar synchronizers, etc. — is a distinct possibility.

UHF TUBES

Interelectrode capacitance and *transit time* are two factors which cause the tubes discussed so far to be inadequate at ultra-high frequencies. Interelectrode capacitance has been treated in the discussion of tetrodes and pentodes. There you noted that the additional grids helped to reduce the feedback caused by interelectrode capacitance. However, as frequencies are increased, the capacitive reactances decrease and sufficient feedback may occur to cause trouble in spite of the additional grids.

Transit time is the time electrons require to travel from cathode to plate. At the lower frequencies, transit time is very much less than the duration of a cycle and is therefore negligible. However, at ultra-high frequencies, the duration of a cycle is so much shorter that transit time is no longer a negligible portion of a cycle. Therefore, transit time causes an additional phase shift between grid voltage and plate voltage, making the phase difference between the two more than 180°. As a result, power output decreases and plate dissipation increases. Transit time, when it is an appreciable portion of the duration of a cycle, also reduces the input impedance of the tube, causing additional loss.

Transit time may be reduced by decreasing the spacing between electrodes or by increasing applied voltages. If the electrodes are moved closer together, the interelectrode capacitances become greater. It is therefore necessary to reduce the size of the elements to keep the interelectrode capacitances small. This reduction in element size and spacing reduces the power handling ability of ultra-high frequency tubes.

An *acorn* tube, shown in the illustration at the bottom of the page, is a miniature tube used at ultra-high frequencies. Its power handling ability is less than that of the tubes used at lower frequencies. Note that the tube has no tube base and that connections are made to the electrodes through pins which protrude from the glass envelope. This keeps leads short and minimizes capacitances between leads. Acorn tubes may be diodes, triodes, or pentodes.

There are physical limitations which determine the spacing and size of the electrodes of miniature tubes. These tubes are therefore not usable at the super-high frequencies. Special tubes such as lighthouse tubes, klystrons, magnetrons, etc., are used at these frequencies. UHF tubes are discussed in greater detail in chapter 34.

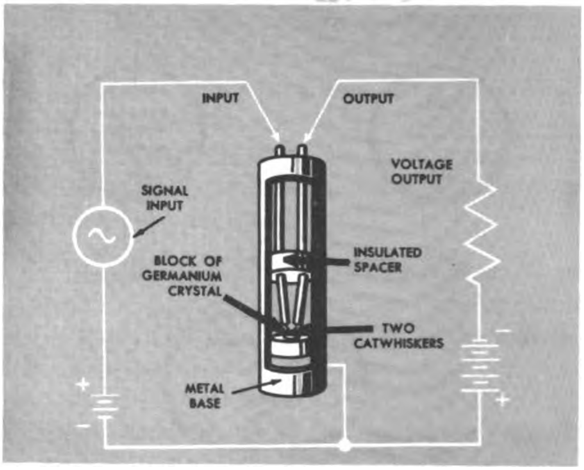
GAS TUBES

Gas-filled tubes have made possible many of the electronic controls of machines in use today. These gas tubes are included in the category of vacuum tubes because they are first evacuated and then filled with gas under low pressure. The gases most frequently used in gas tubes are argon, neon, xenon, and mercury vapor. Gas-filled tubes are often referred to as *soft* tubes in contrast to the *hard* or high vacuum tubes.

In gas tubes, ionization occurs because the gas molecules are bombarded by moving charged particles. Two basic types of gas tubes are used, the *hot-cathode* and *cold-cathode* types. In the hot-cathode type, the cathode is heated in the same way as it is in thermionic tubes; in the cold-cathode type, the cathode is not heated but may become hot as it is bombarded.

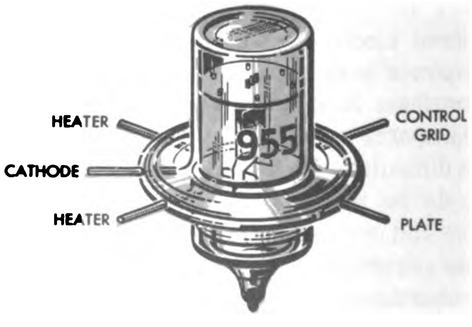
Cold-Cathode Gas Diodes

Cold-cathode, gas-filled diodes are shown schematically in the illustration on page 185. The small dot within the envelope indicates that the tube is gas-filled. The VR tube and the neon glow tube shown have no heaters and are therefore called cold-cathode tubes. Since the cathode is not heated, it does not emit electrons which can ionize the gas. However, a flow of current is detected if a positive potential is applied between plate and cathode.

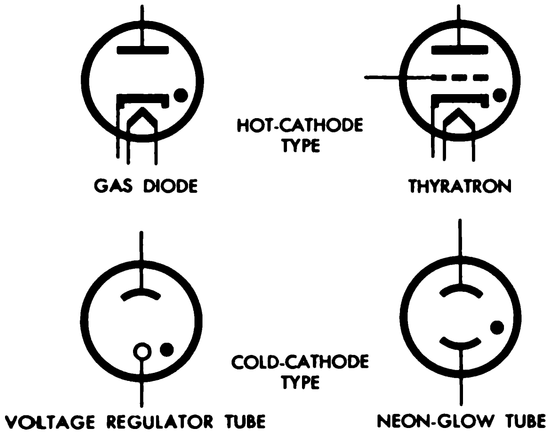


The Transistor

CHARACTERISTIC CURVE. The flow of current may be explained as follows. Some ions and free electrons are always present in gases in their normal state. When the potential is applied across the tube's electrodes, these free electrons and negative ions drift toward the plate, while positive ions move toward the negative cathode. During the movement of the charged particles, numerous collisions take place between the particles and neutral atoms. The collisions result in the freeing of additional electrons and the production of more ions. Since more charged particles are moving, this increases the probability of further collisions and causes the



Acorn Tube



Types of Gas Tubes

production of greater numbers of charged particles. The effect of these charged particles may be noted by measuring the resulting current flow as the potential between the electrodes is varied. Plotting the values of current flow against corresponding values of applied voltage yields a curve similar to the one shown below.

An analysis of the curve gives information about the effect of these moving charged particles. To identify various portions of the curve, several points on the curve are lettered; thus, the part of the curve extending from A to B is REGION A-B.

Region A-B. In this region, the applied voltage begins to cause the drift of ions and electrons within the gas and a small amount of current flows. This current is often referred to as the *dark current* and is of the order of one microampere.

Region B-C. As the voltage is increased, the ionization increases but the ions are neutralized as quickly as they are formed; this accounts for the failure of current to increase over most of this region. However, at some potential near C, the velocity of the charged particles is sufficient to start greater ionization of the gas and current begins to rise sharply. The current flow at point C, the *threshold current*, is on the order of one or two microamperes. The voltage, at which the degree of ionization becomes noticeable, is the *striking potential* or *firing point* of the tube.

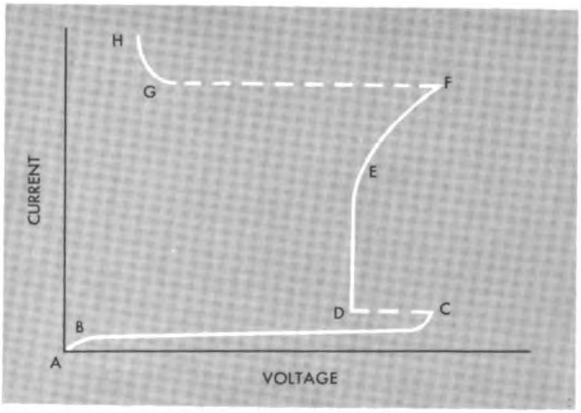
Region C-D. The great increase in the amount of ionization decreases the resistance of the path from cathode to plate. As a result, the voltage drop (I × R) across the tube decreases, as shown by the abrupt decrease in voltage from C to D.

Region D-E. Throughout this region, as the current increases, the voltage drop across the tube remains constant. This is the region most frequently used, and it is referred to as the *operating region*. The reason for the existence of this region is that, as the current increases, the amount of ionization increases; however, as the ionization increases, the resistance of the tube decreases. Therefore, the value of I × R remains constant. The ionization of the gas and the resulting electron flow cause a glow, the color of which is characteristic of the gas used. This type of tube is therefore called a *glow tube;* and the current flow is called a *glow-discharge.*

Region E-F. In this region, a large increase in voltage is required to obtain a noticeable change in current flow.

Region F-G. The current reached at F is so great that an excessive number of positive ions are produced. The comparatively heavy positive ions are attracted to the negative cathode at high velocity (because of the higher voltage), and the resulting bombardment of the cathode is sufficient to heat the cathode to such an extent that it emits electrons thermionically. The addition of these charged particles reduces the resistance of the tube sharply and the voltage drops abruptly to point G.

Region G-H. The increase of electron emission and ionization is cumulative and the current continues to increase until the tube is destroyed. This intense current flow is called an *arc discharge* as compared to the flow discharge in the operating region. The voltage, at which the arc



Characteristic Curve of a Cold-cathode Diode discharge is started (F), is the *breakdown voltage* of the tube.

From the characteristics of cold-cathode gas-filled diodes, you see that, unlike the thermionic vacuum diode, the applied voltage loses control of the current flow once the firing potential is reached. However, a certain potential must be maintained or deionization takes place and current ceases to flow. The potential at which current ceases to flow is known as the *extinction potential*. Also, unlike the thermionic vacuum diode, the voltage drop across the gas tube remains fairly constant with considerable changes in current flow if the tube is operated in the operating region of its characteristic curve.

VOLTAGE REGULATORS. The characteristics of glow tubes are used to advantage in voltage regulation. Some typical tubes used in this manner are the VR-75-30, the VR-105-30, and the VR-150-30. The designation of these tubes indicates the following:

VR—voltage regulator

75,105,150—voltage (within 5%) maintained across tube if current does not exceed 30 ma.

30—Current should not exceed 30 ma. or fall below 5 ma. for good regulation.

Study the characteristic curve and schematic circuit shown in the illustration below, Obtaining the Characteristic Curve of a VR-75-30. The voltage values plotted on the curve are the voltage drops across the tube as read by voltmeter V. The applied voltage E_A is varied by means of the potentiometer P. The resistor R limits the current through the tube and plate current readings are obtained from ammeter A.

From the curve, you may obtain the following information:

The firing potential is 82 volts.

The operating potential is 75 volts.

The extinction potential is 68 volts.

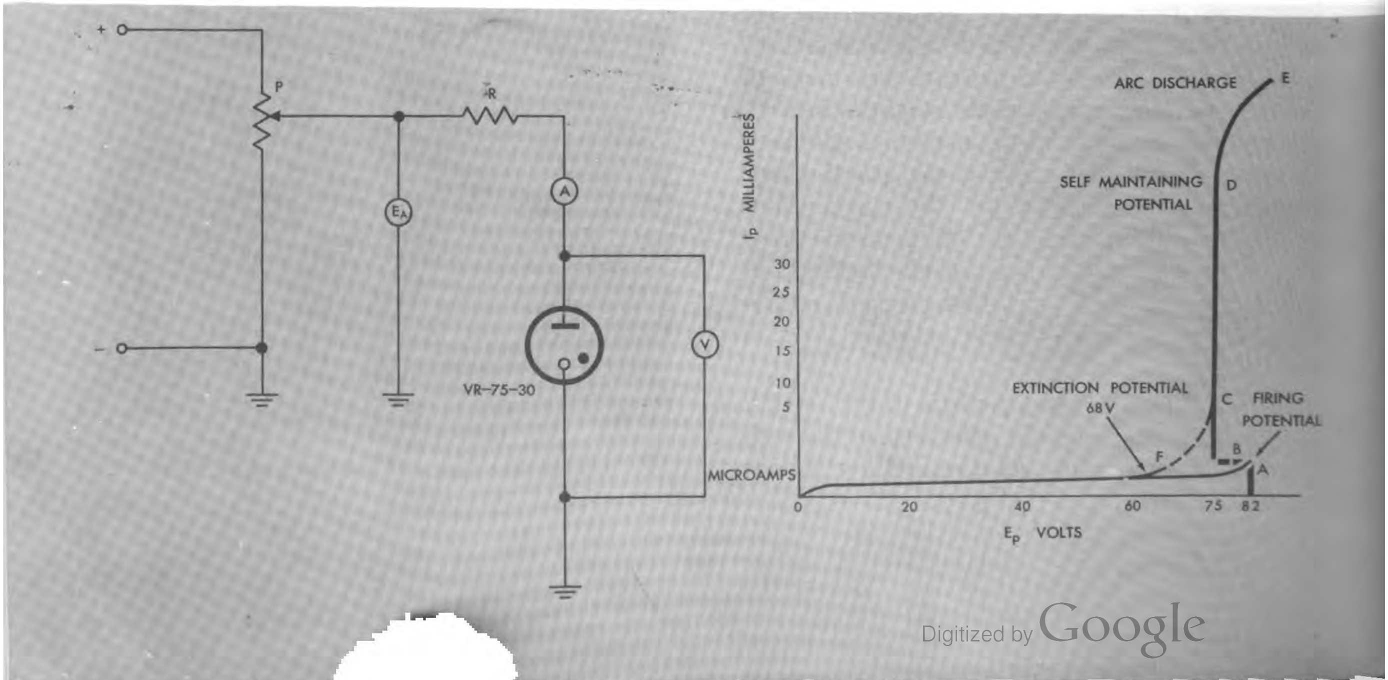
The current rise between points C and D show the working range of the tube.

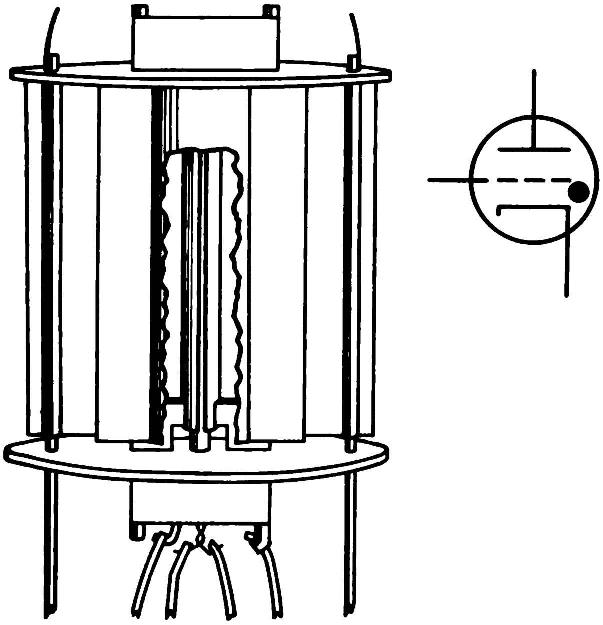
A regulator tube keeps the output voltage relatively constant regardless of whether the *load* or the *input voltage* is varying. If the load current increases, the voltage drop across the limiting resistor (R) increases and reduces the voltage across the tube. A reduction of tube voltage causes a decrease in tube current which, in turn, decreases the voltage drop across R. The voltage across the tube, therefore, remains almost constant. If the input voltage varies, the degree of ionization and the tube resistance change. This maintains a constant voltage across the tube.

Mercury-Vapor Diodes

Where large quantities of rectified current are required, gas-filled tubes with heated cathodes are used in preference to thermionic vacuum diodes. The gas used in such tubes may be any of those previously mentioned; however, the *mercury-vapor diode* is the most popular. Although high vacuum diodes draw currents of less than one ampere, some mercury-vapor

Characteristic Curve of a VR-75-30





Construction of a Thyratron

diodes are capable of handling several hundred amperes of current. The ability of mercury-vapor tubes to pass large quantities of current is accounted for, to some extent, by the low voltage drop across the tube. Since the power dissipated by a tube is EI, the amount of power dissipated for a given value of current is proportional to the voltage drop across the tube. In gas tubes where the voltage drop is smaller than it is for vacuum diodes, more current may flow before the power dissipated by the tubes becomes excessive and the tube overheated. The mercury-vapor diode is therefore used in rectifier circuits which require high current, low voltage, and good regulation.

THYRATRONS

The addition of a control grid to a hot-cathode gas diode yields a gas-filled triode known as a *thyratron* (door) tube. Thyratrons have many important applications in the industrial electronics field, as well as in the fields of radio, radar, and television.

Construction

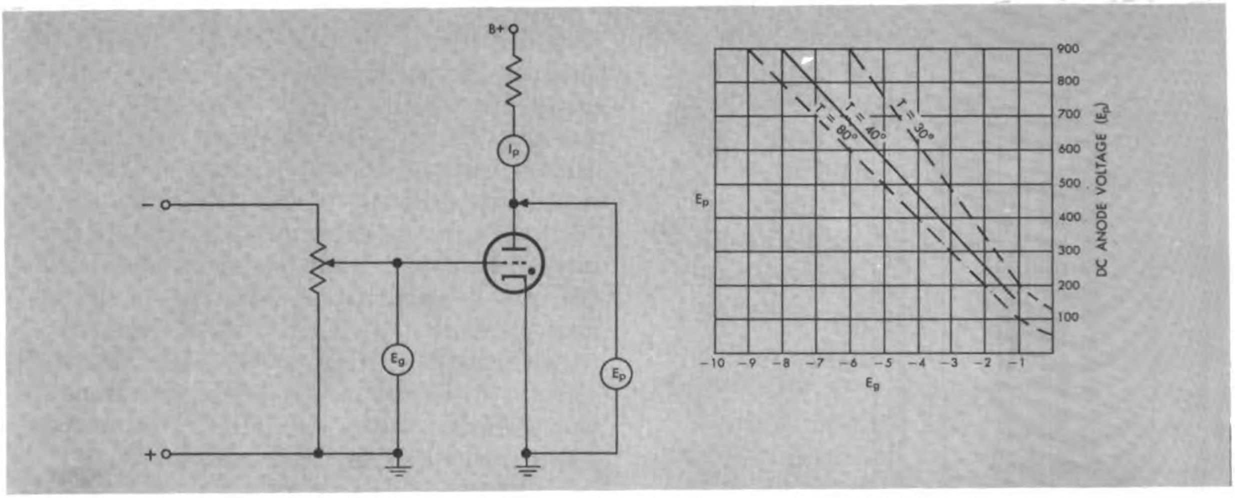
The construction and schematic representation of a typical thyratron (884) is shown in the illustration above. Note that the grid structure is rather large and that it consists of a metal cylinder with a tiny vertical slit. The control grid acts as a shield between the cathode and the plate. The size of the slot in the grid and the potential on the grid determine how many of the electrons emitted by the cathode enter the space between the grid and plate. The number of electrons which enter this space and the plate potential determine the degree of ionization of the gas. A small negative voltage on the grid can therefore suppress the effects of a much larger positive voltage on the plate. Although the gas in the 884 tube is argon, thyratrons are manufactured which are filled with mercury vapor, helium, or any other inert gas.

Operating Characteristics

When the voltage between the grid and cathode is of such value as to permit enough electrons to pass through the grid openings to cause ionization of the gas (at the given plate voltage), current flows through the tube in large quantities. The value of the current flow depends only on the impedance of the external circuit, the grid having lost practically all its control. This loss of control by the grid may be explained as follows:

The collisions between the thermionically emitted electrons and the atoms of the gas in the tube cause the formation of positive ions. These are attracted by the negative grid and form a sheath around it, neutralizing the effects of the grid. If the grid potential is made more negative, more of the positive ions are attracted to it, and the grid remains neutral. The triode then behaves like the hot-cathode diode in which positive ions neutralize the effects of the space charge. This results in large currents with low voltage drops across the tube.

Before the grid can again control the flow of current through the tube, the plate voltage must be reduced to a value below the ionization potential of the gas. However, it takes time for the ions and electrons to recombine to form neutral atoms, and for the ion sheath around the grid to be removed. The *deionization time* varies with the type of gas, electrode voltages, pressure, and temperature. For most applications, this deionization time is so short (measured in microseconds) that it may be neglected. However, at extremely high frequencies, the deionization time makes the use of gas tubes impracticable.



Effect of Grid Voltage on Firing Point of Thyratron

Grid Voltage and Firing Point

The firing point for a given thyratron depends upon the grid voltage — the more negative the grid, the higher the plate potential necessary to cause the tube to fire. You may use the circuit illustrated above to show the effect of grid voltage on the firing points of a given thyratron.

Vary the plate voltage (E_p) until the tube fires for a given value of grid voltage (E_g). Change the grid voltage and repeat the process, and another value of firing potential is obtained. Continue this process for various values of grid voltage and plot the results. The curve plotted shows the firing potential of the tube for various values of grid voltage.

Three curves which show the operating characteristics of a thyratron are drawn in the illustration, Operating Characteristics of a Thyratron. The points which lie to the right of the line represent conduction; while those points that lie to the left of the line represent nonconduction. If the temperature is changed, the firing point changes; the curve shifts to the left with an increase in temperature and to the right with a decrease in temperature. These curves indicate that, with the same grid voltage, a thyratron fires at a lower plate potential if the temperature of the gas is higher. This is due to the fact that the atoms of gas are more agitated at the higher temperatures and are therefore more easily ionized.

Since the curves shown are straight lines for the most part, a definite ratio exists between the change in grid voltage (E_g) and the change in firing potential (E_p). This ratio, called the *control ratio*, is defined as the increase in plate voltage necessary to cause the tube to fire if the grid voltage is made one volt more negative. At 40°, the tube whose curves are plotted has a control ratio of 110. This may be found by noting that a 2-volt grid change (from −2 volts to −4 volts) requires an increase of 220 volts on the plate (from 250 volts to 470 volts). This is equivalent to a ratio of 110:1 or 110. The control ratio of any thyratron may be found in similar manner. For example, a thyratron which fires at 80 volts with a grid voltage of −6 volts and at 90 volts with a grid voltage of −7 volts has a control ratio of 10:1 or 10.

USES OF THYRATRONS

The thyratron has many useful applications in industry. It is widely used as an electronic switch in counting and control circuits. In radar applications, thyratrons are used in driver circuits. When used in this way, the tube is normally cut off by a negative grid voltage. It is made to conduct by a positive trigger pulse applied to the grid. Once the tube fires it continues to conduct until the plate voltage falls below the extinction potential. The negative grid voltage then regains control and keeps the tube from conducting until the positive trigger pulse causes it to fire again.

Thyratrons are also used as sawtooth generators. The manner in which a thyratron is used to generate a sawtooth sweep voltage is discussed on page 314.

Chapter 23

Cathode-ray Tubes

The cathode-ray tube (CRT) is a vacuum tube especially designed to display electrical signals so that they may be examined or measured. You are already familiar to some extent with the oscilloscope which consists of a cathode-ray tube and its associated circuits. Cathode-ray tubes are used in radar equipment to display visually the information received. They perform a similar function with the picture portion of the received television signal. The transmission of the picture portion of the television signal is made possible by the use of special types of cathode-ray tubes. These and many other uses for cathode-ray tubes are found in the field of electronics.

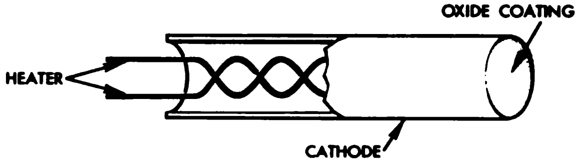
All cathode-ray tubes consist of an electron gun that projects a beam of electrons along the axis of the tube, a means of deflecting and focusing the beam, a fluorescent screen which changes the energy contained in the beam into light, and an evacuated glass envelope. The beam of electrons may be deflected and focused by an electric field or a magnetic field. An *electrostatic CRT* uses electric fields for focusing and deflecting the beam. An *electromagnetic CRT*, on the other hand, focuses and deflects the beam by means of electromagnetic fields.

THE ELECTROSTATIC CRT

The vacuum tubes studied thus far all have one common feature — a centralized cathode surrounded by various electrodes. Electrons emitted from the cathode are sent radially in all directions. In the CRT, however, the electrons emitted from the cathode are confined to a parallel beam by means of an *electron gun.*

The Electron Gun

The electron gun of the electrostatic CRT consists of a cathode which is the source of electrons, a grid which concentrates the electrons into a beam, and anodes which focus the beam of electrons on the screen. These are described in the following paragraphs.



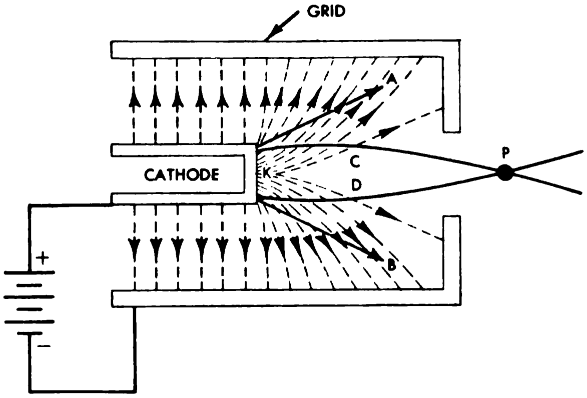
Cathode of a Cathode-Ray Tube

THE CATHODE. The CRT has an indirectly heated cathode which consists of a small metal thimble about ⅛ inch in diameter within which a hot wire is mounted. To make this thimble a good emitter when heated, the front end is concave and oxide-coated. A partly cutaway illustration of a cathode is shown above. Note that emission is out of the front end only.

The tungsten wire heater coil is wound in a double spiral so that the electron beam will not be affected when AC is applied. In this way, the magnetic field of each winding is canceled by the equal and opposite field of the other. The heater coil is insulated from the cathode; however, the insulation touches the cathode to provide good heat conduction. Heater voltage for the CRT is usually 6.3 volts, AC.

THE GRID. The grid, a metal cylinder having a metal diaphragm with a small hole, surrounds the cathode. The potential applied to the grid is negative with respect to the cathode; therefore the electrostatic field between the two is as shown in the illustration below.

An electron which starts out in the direction KA is bent by the lines of force so that it takes the path KCP. In similar fashion, the electron which starts out in direction KB is bent so that it takes path KDP. The paths of these two electrons intersect at P. All electrons which pass through the hole of the grid are concentrated in this manner into a beam along the axis of the tube.



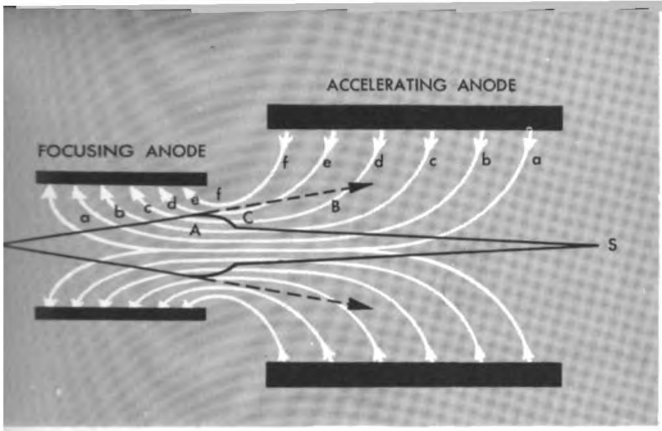
The lines of force between grid and cathode prevent the emitted electrons from diverging outwards away from the cathode and cause them to be bent inward along the tube's axis. They may even cause the electrons to cross over and diverge on opposite sides of the axis. The potential between grid and cathode determines how many electrons are brought along the axis and how many diverge and are lost inside the electron gun structure. The grid of the CRT therefore functions in two ways: first, it controls the amount of space current flowing from the cathode; second, it acts as a lens which concentrates the electrons into a beam.

The intensity of the spot on the fluorescent screen depends on the number of electrons in the beam; the number of these electrons is controlled, in turn, by the voltage between grid and cathode. A potentiometer which varies the bias on the grid is used with most cathode-ray tubes to enable an operator to cause the spot to become brighter or dimmer. This potentiometer is known as an *intensity*, *brilliance*, or *brightness* control.

THE ANODES. The electron gun in the electrostatic CRT consists of two anodes in addition to the cathode and grid. These anodes are cylinders which have metal diaphragms with holes in them. The *first anode*, sometimes called the *focusing anode*, is closer to the grid and is positive (a few hundred volts) with respect to the cathode. Behind the first anode is another cylinder of larger diameter called the *second* or *accelerating anode*. This anode is at a much higher positive potential (several thousand volts in some cases).

As you have seen in the previous illustration, the grid focuses the electrons into a beam at point P which is close to the grid. However, the electron beam must travel to the screen and be focused there. The electrostatic field between the two anodes makes this possible. The lines of force between the two anodes are shown in the illustration at the top of page 191.

An electron introduced in the field at point P would tend to follow one of the lines of force and hit the more positive anode. Because of the posi-



Lines of Force between Two Anodes of a Circuit

tive voltage on the anodes, however, the electron travels at a very high velocity and does not stay in the field long enough to be pulled completely off its course. But it is subjected to forces as it moves through the field. For example, an electron which is traveling from point P along the path PAB is curved toward the axis by the lines of force and is forced to take path PAC.

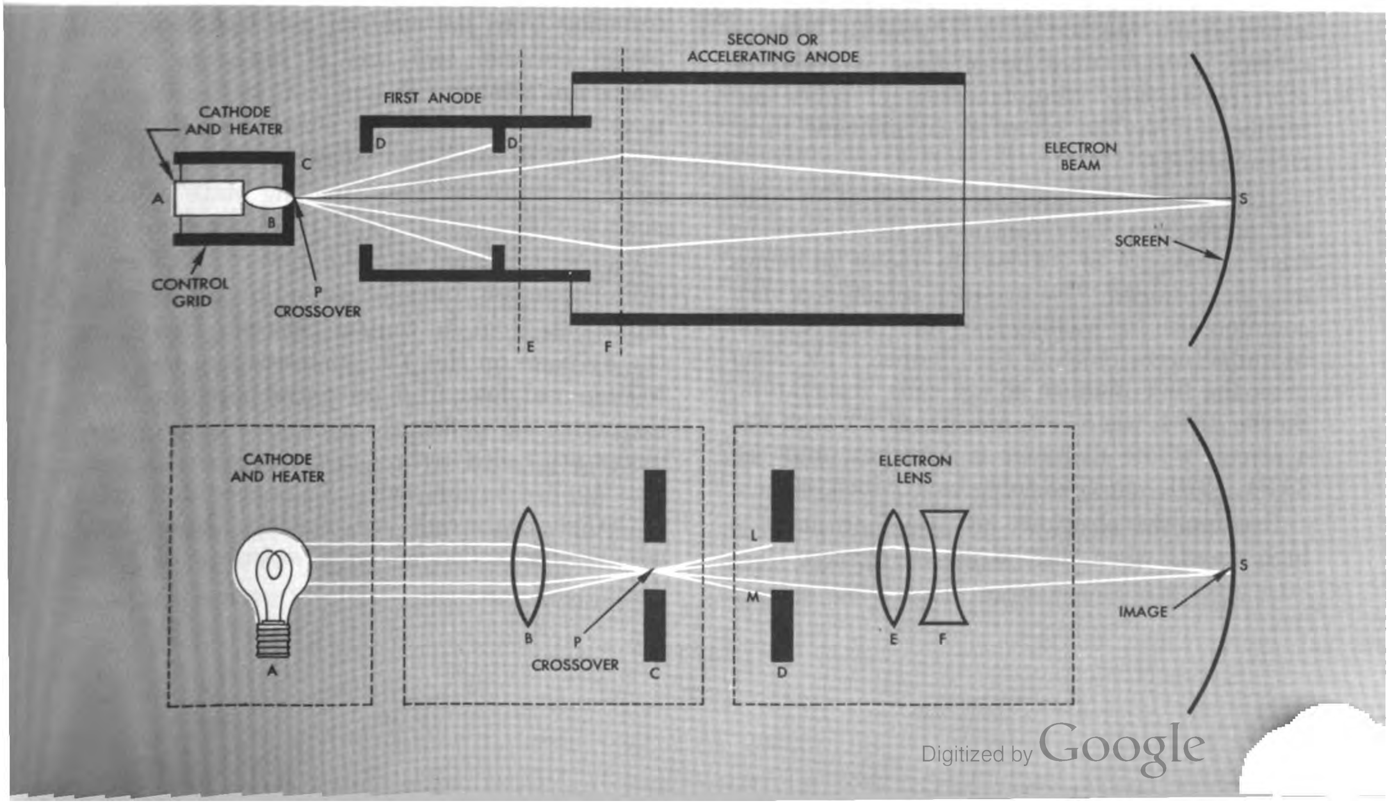
The lines of force near the right end of the second anode are almost parallel to the axis of the tube; therefore, the beam is accelerated in that direction and moves toward point S. The forces tending to push the electrons toward the axis is greater for those which are farther from the axis than for those which are nearer. All electrons therefore converge at point S which is on the screen. The point where the electrons converge may be shifted by varying the strength of the electric field between the two anodes. A potentiometer which varies the voltage on the first anode is usually used for this purpose. This control which focuses the beam on the screen is called a *focus* control.

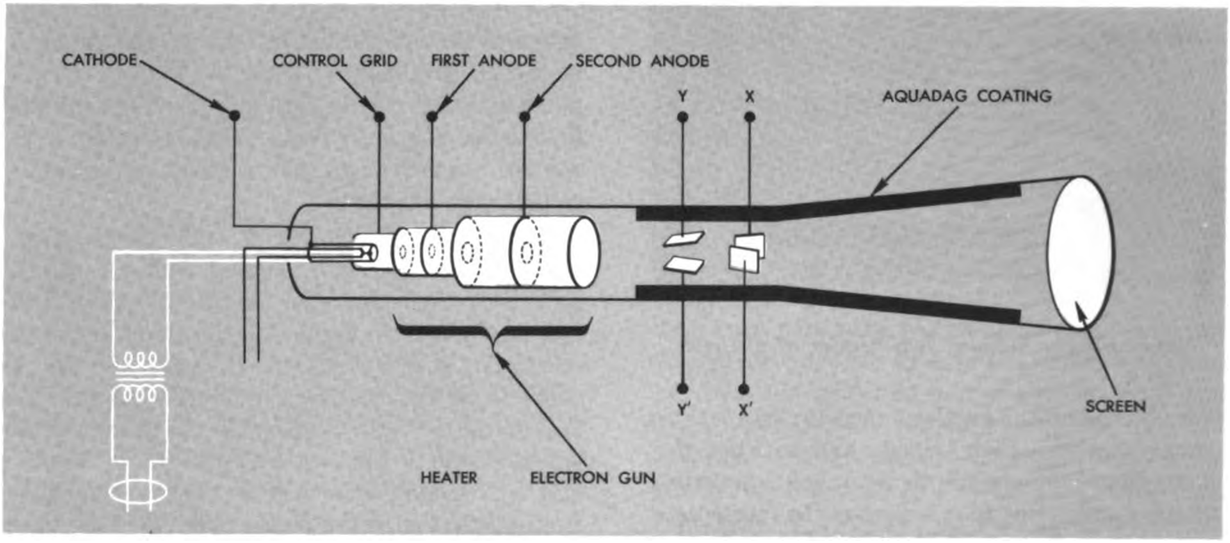
A complete electron gun and screen of an electrostatic CRT is compared with an optical system in the illustration below. In the optical system, light from lamp A is caused to converge to a point at P by the lens B. Light appears to come from the crossover point P for the lens system which follows. A proper adjustment of lenses E and F focuses the light on the screen. LM is a second lens used to prevent the light from scattering in the optical system. The same letters are used in both systems to indicate parts that perform corresponding functions. Electron lenses function in a manner similar to light lenses; however, they have the advantage of having an adjustable focal length which is readily controlled by varying a voltage.

A Complete Electrostatic CRT

In addition to the electron gun, the electrostatic CRT consists of a highly evacuated glass tube, two sets of deflection plates and a fluorescent screen. This is shown in the illustration, An Electrostatic CRT on page 192.

Comparison of Electron and Optical Lens System





An Electrostatic CRT

THE SCREEN. The fluorescent screen is covered with a coating of chemicals deposited on the inside walls of the end of the tube. These chemicals or phosphors which are rendered active by adding traces of other substances may be one of the following: willemite, calcium tungstate, phosphorescent zinc sulphide, etc. When fast-moving electrons strike a screen coated with one of these chemicals, visible light is emitted. The color of this visible light — green, white, yellow, or blue — depends upon the chemical used. After the electron impact, the emission of light persists for a small fraction of a second. This *persistence* depends upon the nature of the screen coating and serves as one method of classifying the various screens; screens may be of *long*, *medium*, or *short* persistence. Manufacturers of tubes use numbers after the letter P to indicate these classifications.

Screens differ in diameter as well as in color and persistence. Variations in the sizes of screens from a diameter of 2 inches to one of 28 inches may be obtained. The five-inch screen is a very common size for oscilloscopes and radar scopes. Manufacturers indicate this diameter by using it as the first number of the tube designation. So the 5FP4, the 5TP4, and the 5CP1 each has a screen whose diameter is 5 inches; and the 2AP1, 2BP1, and 2P23 tubes have 2-inch screens.

Because the electrons are traveling at very high speeds when they strike the screen, they cause secondary emission from the screen. These secondary electrons, if allowed to remain would form a space charge around the screen and interfere with normal tube action. To prevent this space charge from accumulating, the inside of the CRT is coated with graphite. The *aquadag*, as this graphite coating is called, extends to about one-half inch from the screen. Since it is at a positive potential with respect to the cathode, the aquadag removes the secondary electrons and also serves as a shield for the electron beam.

DEFLECTION PLATES. The electron beam, in traveling from the electron gun to the screen, passes between two sets of *deflection plates*. These are labeled YY′ and XX′ in the illustration. When one plate (of either pair) is made positive with respect to the other, it attracts the beam of electrons. Thus, making plate Y more positive than plate Y′ (or Y′ more negative than Y) causes the beam to move upward. As viewed from the screen, the spot caused by the beam is moved up from its central position. Reversing the voltages on the YY′ plates causes the spot to be moved down from its central position. Since the plates YY′ deflect the spot up or down, these plates are called *vertical deflection* plates. Voltages applied to the plates XX′ cause deflection of the spot to the right or to the left. These plates are therefore called *horizontal deflection plates*.

Sensitivity of Deflection Plates

The amount of deflection caused by a pair of deflection plates depends upon the velocity of the electron beam, the voltage applied between the plates, and the spacing between the plates. *Deflection sensitivity* is defined as the number of millimeters of movement of the spot on the screen if one volt DC is applied between the plates. Thus, a deflection sensitivity of .220 mm/volt DC means that each volt applied between the deflection plates causes a deflection of .22 millimeters.

A more useful measurement of the sensitivity of deflection plates is known as the *deflection factor* — measured in volts per millimeter or volts per inch. This factor indicates the voltage that must be applied between a pair of deflection plates to cause a deflection of the spot on the screen of one millimeter or one inch. Tube manuals give the deflection factor in volts per inch.

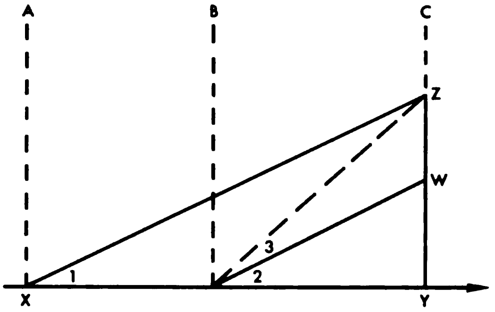
The deflection sensitivity and deflection factor are reciprocals; that is, $\mathrm{ds} = \frac{1}{\mathrm{df}}$ and $\mathrm{df} = \frac{1}{\mathrm{ds}}$.

The deflection factor of the vertical plates differs from the deflection factor of the horizontal plates because they are at different distances from the screen. For example: A CRT in which the spot moves ½ inch upward if 50 volts DC is applied between the vertical deflection plates and ¼ inch to the right if 30 volts DC is applied between the horizontal deflection plates has the following deflection factors.

$$\text{vertical df} = \frac{50}{.5} = 100 \text{ volts per inch}$$

$$\text{horizontal df} = \frac{30}{.25} = 120 \text{ volts per inch}$$

In this case the vertical deflection plates are more sensitive than the horizontal deflection plates; that is, they require less voltage for the same amount of deflection of the spot on the screen. The vertical deflection plates are more sensitive in most applications because they are farther from the screen than the horizontal deflection plates. The illustration on this page shows why this is true. The position of the vertical deflection plates is represented by A, of the horizontal deflection plates by B, and of the screen by C. If a voltage applied at A causes the beam to be deflected at angle 1 from its normal path of travel (XY), the spot on the screen is displaced a distance YZ. The same voltage applied at B would cause the beam to be deflected at angle 2 (equal to angle 1) and the spot to be displaced at distance YW. To cause the spot to be displaced a distance YZ on the screen requires a voltage applied at B great enough to cause a deflection angle of 2+3.



Effect on Deflection of Distance of Deflection Plates from Screen

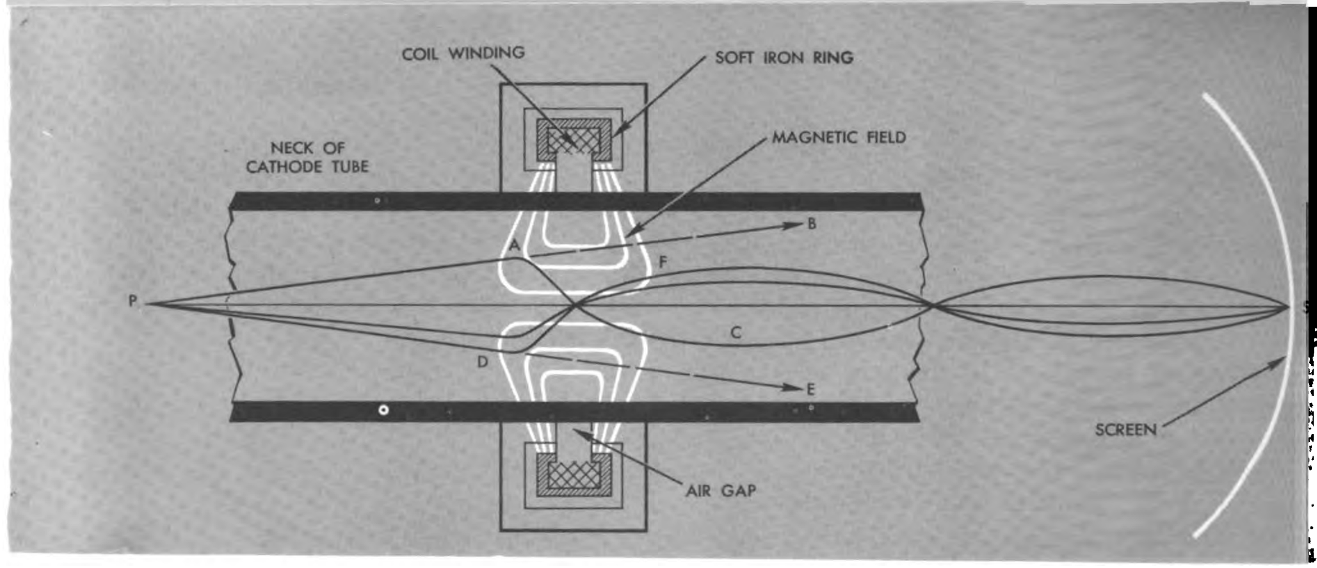
Although cathode-ray tubes are usually mounted so that the vertical plates are the more sensitive ones, situations sometimes arise where it is desirable to make the more sensitive plates cause horizontal deflection. The RCA Tube Handbook therefore labels the deflection plates DJ_1, DJ_2, DJ_3, and DJ_4. DJ_1 and DJ_2 are the deflection plates nearer the screen — usually horizontal deflection plates. DJ_3 and DJ_4 are the deflection plates farther from the screen and are therefore more sensitive — usually the vertical deflection plates.

THE ELECTROMAGNETIC CRT

The filament, cathode, and control grid of the electromagnetic type of CRT are similar to those of the electrostatic type. However, focusing and deflection of the beam are accomplished electromagnetically; that is, by means of magnetic fields. Although this type of CRT also has two anodes, (here, the Aquadag coating is the second anode), these anodes serve the purpose only of accelerating the electrons and attracting them to the screen with great velocity. Focusing of the beam may be accomplished by means of a coil or a permanent magnet.

The Focusing Coil

The illustration, Electromagnetic Focusing, shows a cross section of the magnetic field set up inside the neck of a cathode ray tube by a focusing coil wound inside a soft-iron ring. Al-



Electromagnetic Focusing

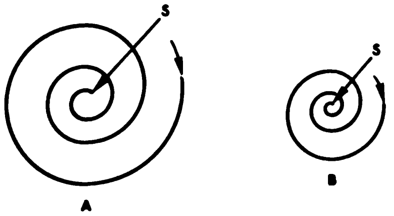
though the coil may be wound without any iron near it, the presence of the iron makes it possible to develop a magnetic field of the necessary strength with less current. The iron also serves the purpose of concentrating the field strength where it is needed. An air gap makes it possible to concentrate the field within the cathode-ray tube. A direct current which flows through the coil windings sets up the magnetic field. The amount of current through the coil — hence, the magnetic field strength — may be varied by a rheostat, the *focus control.*

The manner in which a beam is focused may be explained with the aid of the illustration above. Point P is the crossover point which the electron beam is assumed to have reached after being emitted by the cathode, accelerated by the first anode, and acted upon by the grid in the same manner as for the electrostatic CRT. An electron traveling from point P along path PAB and entering the field of the focus coil is traveling in a direction which may be resolved into two components: from *left to right* and *upward.* One component is parallel to the axis and the magnetic field; the other is at right angles to the axis and the magnetic field.

Consider each component separately. An electron being pulled from left to right is not affected by the parallel magnetic lines of force. It therefore continues along the path in which it is being pulled. An electron being pulled upward, however, is affected by the perpendicular magnetic lines of force. This electron (or beam of electrons) may be considered to be similar to a conductor in a perpendicular or magnetic field; it is forced in a direction which is at right angles to the field and to the direction in which it is moving. In this case, the electron would be forced at right angles to the page. As the electron changes direction, the magnetic field continues to act upon it and causes it to move in a circular path around the axis. Thus the two components combine to cause the electron to continue moving from left to right and, at the same time, travel in a circular path around the axis. The resultant effect of the magnetic field on the motion of the electron is to make it travel a helical path (like the thread on a wood screw) while passing through the magnetic field. Thus the electron shot into the field along path PAB is forced to take the spiral path PACS to the screen.

In similar fashion, the electron shot into the field along path PDE is forced to take the spiral path PDFS. All electrons coming from the crossover point P and entering the field of the magnet follow paths that make them meet at a common point S. The spiral path is greater for the electron entering the magnetic field at a greater angle; however, they travel faster and reach point S at the same time as the electrons entering the magnetic field at a smaller angle.

If you could look down on the face of the scope and see the paths of the electrons, you would see spirals like those shown in the next illustration. A is the path of an electron which enters the magnetic field at a greater angle than B. Each spiral path has the same number of turns and all electrons arrive at S at the same time. In order to focus the beam, it is only necessary to adjust the strength of the magnetic field so that the common point S falls on the fluorescent screen. This may be done by varying the current through the focusing coil.



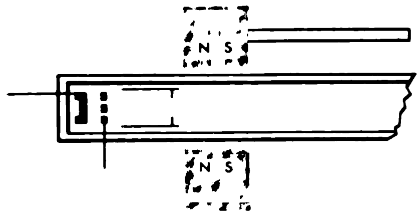
Paths of Electrons toward Screen

In some cases a permanent magnet in the form of a ring is placed around the neck of the CRT. Focusing is accomplished by moving the ring along the axis of the tube. Usually, however, the magnet is a two-section ring, as shown in the illustration below. Varying the size of the air gap focuses the beam.

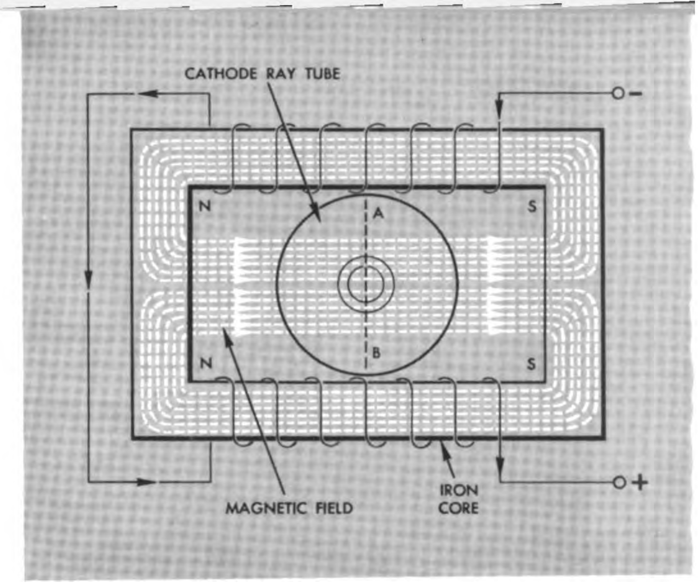
Deflection Coils

Deflection in the electromagnetic CRT, like focusing, is accomplished by means of a magnetic field set up by current flowing through coils. Deflection in one direction may be obtained by winding coils around the neck of the tube as shown in the illustration in the next column.

Here the two coils are in series and wound on an iron core so that the field produced by one coil opposes that of the other. The magnetic lines of force pass through the neck of the tube to complete the magnetic circuit. If the electron beam is assumed to be coming out of the page toward the screen, the magnetic field shown will deflect the beam downward. The deflection is downward because the magnetic field around the electron beam is clockwise. Therefore, it is in the same direction as the magnetic field of the deflection coil at the top and in the opposite direction at the bottom. If the current through the coil is reversed, the direction of the magnetic field is reversed and the beam is deflected upward. With the coils stationary in the position shown, deflection can take place only along diameter AB.



Permanent Magnet for Focusing Electron Beam



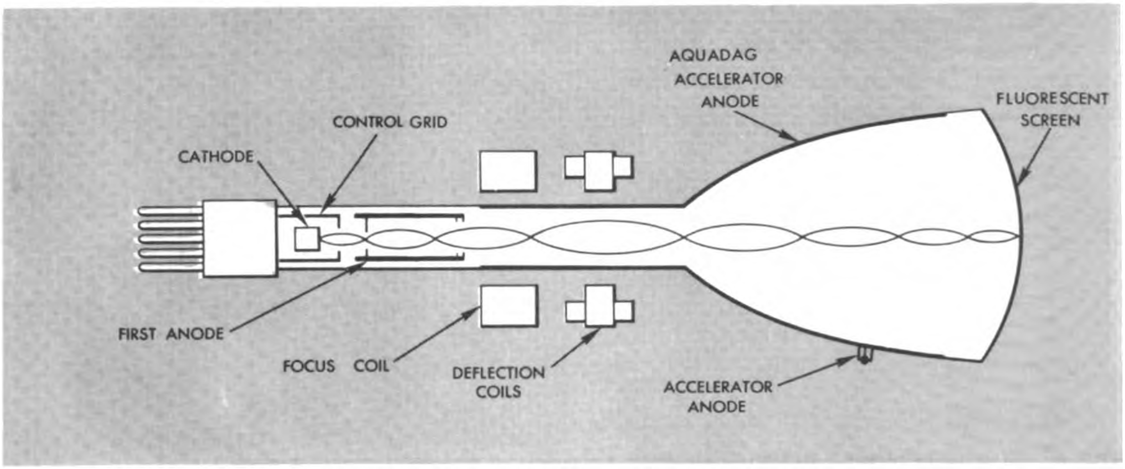
Deflection Coil for an Electromagnetic CRT

It is possible to rotate the entire yoke around the neck of the CRT and produce deflection along any diameter. If rotation is continuous, a constantly rotating sweep may be produced. In radar sets, this type of rotating sweep is produced for a PPI (Plan Position Indicator) scan.

A Complete Electromagnetic CRT

A complete electromagnetic CRT is shown in the illustration on page 196. The shape of this CRT differs somewhat from that of the electrostatic type in that the neck of the tube is very narrow. This is done so that the focusing and deflection coils may be mounted very close to the electron beam. The glass tube expands very rapidly from the narrow neck; electrons may therefore be deflected over a large screen without striking the walls of the tube.

Note the position of the *second* or *final* anode. This anode is at a very high potential (about 5000 volts for a 5-inch CRT) with respect to the cathode. It consists of an aquadag coating extending over much of the inner surface of the tube. The aquadag performs the same function as it does in the electrostatic CRT—prevents a space charge from forming by removing secondary electrons. This final anode is sometimes called the *accelerating* anode because its high potential accelerates the emitted electrons so that they are traveling at considerable velocity when they strike the screen.



An Electromagnetic CRT

The intensity or brightness of the spot on the screen of this CRT is controlled in the same way as it is in the electrostatic type — by varying the bias on the grid. Care must be taken to keep the grid from going positive or the beam current will damage the screen.

Focusing is usually accomplished, as already explained, by varying the current through the focusing coil. This differs from the method used in the electrostatic type in which the voltage on the focusing anode is varied.

The signal to be observed is applied between the grid and cathode in the electromagnetic type of CRT. A signal therefore varies the brilliance of the spot and is indicated by a brightening of the screen. This contrasts with the electrostatic type of CRT in which the signal is applied between a pair of deflection plates and is indicated as an amplitude change on the screen.

A COMPARISON OF CRT'S

Both types of CRT's have advantages and disadvantages. The electrostatic type requires very little current (or power) to accomplish deflection of the beam. The auxiliary circuits of the electrostatic type are much simpler. This type is therefore used in most oscilloscopes. The electromagnetic type has the advantage of being simpler in construction since it has no deflection plates or focusing anode within the glass envelope. It has greater ruggedness — an advantage in mobile equipment — and a shorter tube length which allows for a reduction in overall size of equipment. Its main disadvantage is that it requires a great deal of power to accomplish deflection of the beam.

A summary and comparison of the two types of CRT's as used in conventional circuits is presented below.

ITEM COMPARED	ELECTROSTATIC CRT	ELECTROMAGNETIC CRT
Use	Oscilloscopes, some radar indicators, small television sets, etc.	Radar indicators, large television sets, etc.
Type of Deflection	Electrostatic — using plates	Electromagnetic — using coils
Sweep Applied As	Voltage on deflection plate	Current through deflection coils
How Focused	By varying the voltage on the 1st Anode	By varying the current through a focusing coil or permanent magnet
Intensity or Brightness Control	Varies bias voltage	Varies bias voltage
Signal Usually Applied	To deflection plate	To control grid
Second Anode	Part of electron gun; usually at ground potential	Aquadag — at very high positive potential
Centering Controls	Vary voltage on deflection plates	Vary current through deflection coils
Signal Appears As	A picture of the voltage wave	A variation in the intensity of the beam
Current Required for Deflection	Very low	Very great
Filament Voltage	Usually 6.3 volts	Usually 6.3 volts

Chapter 24

Introduction to Radio Transmission and Reception

A system whose purpose it is to transmit information from one point to another is called a *communication* system. Radio is an instantaneous communication system in which information is sent from one point through space and received at another point. It is the only practical means for communicating with moving vehicles such as boats, airplanes, and motor-driven conveyances.

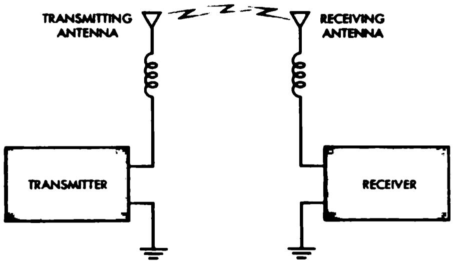
A radio communication system, as you can see in the illustration at top of page 198, consists of three principal parts: a *transmitter*, an *antenna system*, and a *receiver*. The transmitter generates radio-frequency power and modulates it with the information to be sent. Transmission lines carry this radio-frequency power from the transmitter to the transmitting antenna which radiates energy into space in the form of electromagnetic radiations. A receiving antenna in the path of these radiations picks up some of this radiated energy. Transmission lines carry this received power to the receiver, which converts the power into a form from which the information can be extracted.

TRANSMITTER

A radio transmitter consists of two sections. (See block diagram, page 198.) One section generates radio-frequency waves and amplifies these waves so that they have sufficient output power, and the other modulates the radio-frequency waves with the desired information to be transmitted.

Radio-Frequency Section

The radio-frequency section of a transmitter consists of an oscillator and amplifiers. The oscillator is an electronic circuit which generates



Block Diagram of a Radio Communication System

radio frequency energy at the desired frequency. The radio-frequency which is generated by the oscillator is called the *carrier* frequency because it is used to carry the information through space. Many different types of oscillators are suitable for generating the carrier frequency. In chapter 27, you will find a more complete discussion of the more commonly used types.

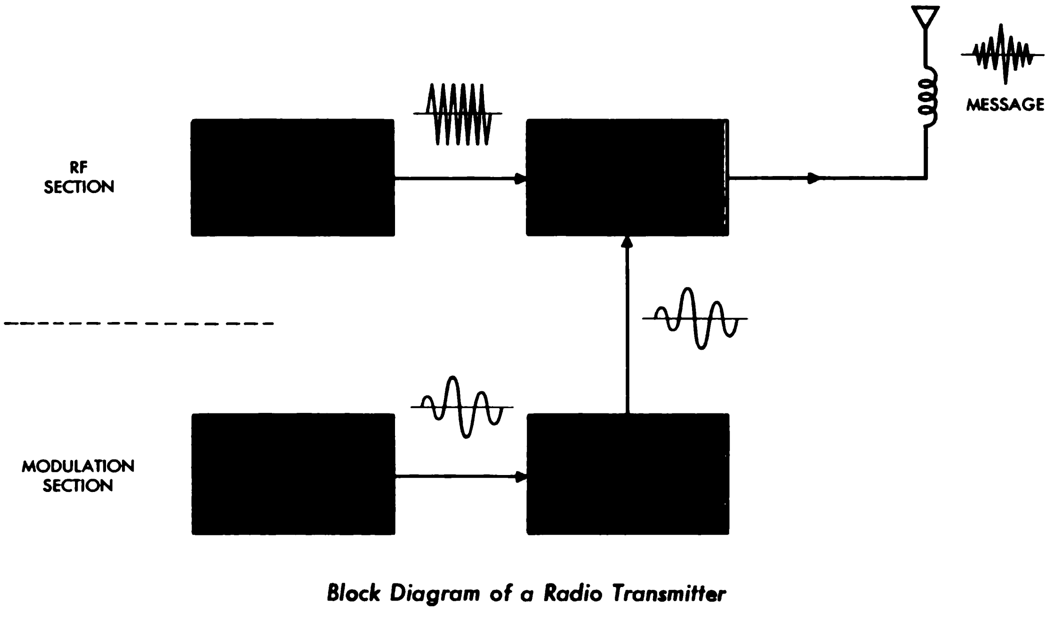
If the electromagnetic radiations are to carry the information any distance, they must have sufficient power. The radio-frequency energy generated by the oscillator is, in most cases, very small. This power must, therefore, be increased if the carrier wave is to serve its purpose. Amplifiers are used to increase the power output of the transmitter. These amplifiers are known as radio-frequency amplifiers.

Modulator Section

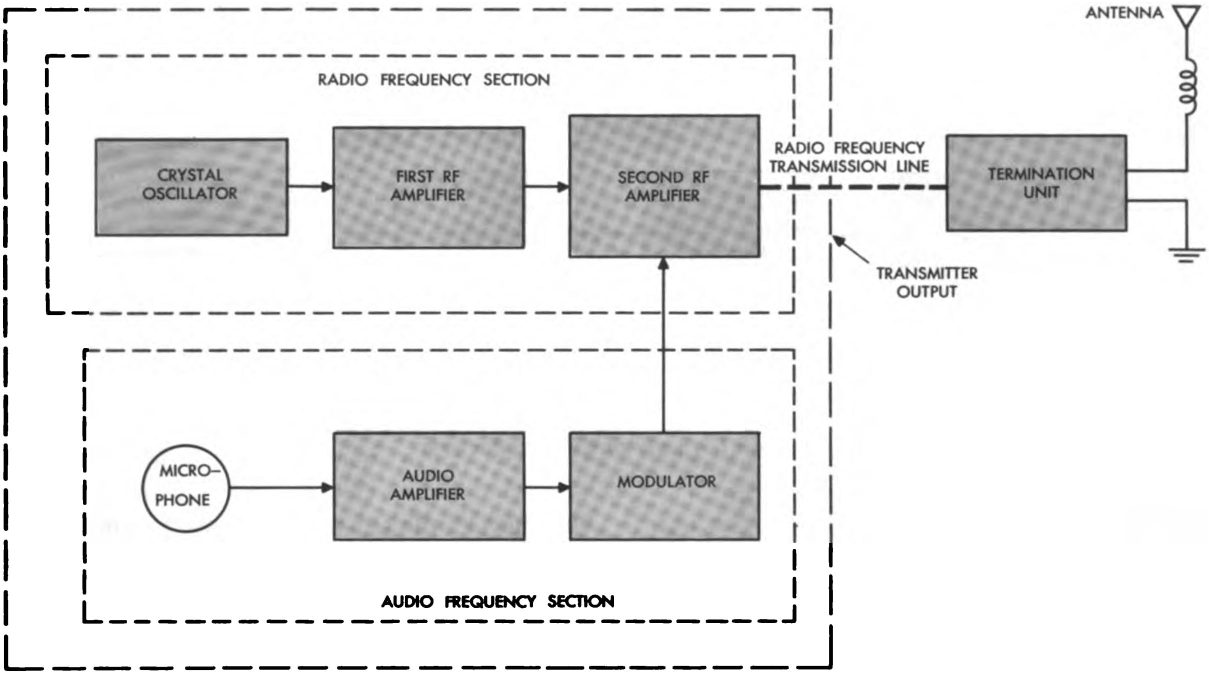
To send information over any distance, it is necessary to vary some steady signal in accordance with the information. The American Indian did this even in his crude method of sending information by means of smoke signals. If the steady smoke signals were sent without variation, no information was communicated. However, if the rising smoke was interrupted so that puffs of smoke resulted, information was carried to an observer some distance away.

A radio broadcasting station which transmits an unvarying carrier signal sends no information to the radio receivers that are tuned in. The carrier must be changed in some way — in amplitude, in frequency, or in phase — to serve as a means of imparting information. This process of varying a carrier so that it can convey a message is called *modulation*. The message is called the *modulating signal*, and the changed carrier is called the *modulated wave*. Amplitude modulation (AM), frequency modulation (FM), and phase modulation refer, respectively, to changes in amplitude, frequency, or phase of the carrier wave. The modulator section of the transmitter, shown in the illustration below causes amplitude changes in the output of the radio-frequency section in accordance with the information being sent. The modulated carrier is therefore an AM wave.

A block diagram of a typical radiotelephone transmitter is shown in the illustration on the next page. This transmitter has a radio-frequency section and an audio-frequency section. The carrier wave is generated by a crystal-



Block Diagram of a Radio Transmitter



Block Diagram of a Radiotelephone Transmitter

controlled oscillator and amplified in two radio-frequency amplifiers. The information to be transmitted is an audio-frequency signal picked up by a microphone. It may be a voice message, music, or any type of sound signal. The energy produced by the microphone is very weak, so it is amplified in two stages of amplification. The second audio amplifier stage is the modulator, because it causes the amplitude of the RF carrier wave to vary in accordance with the audio frequency signal.

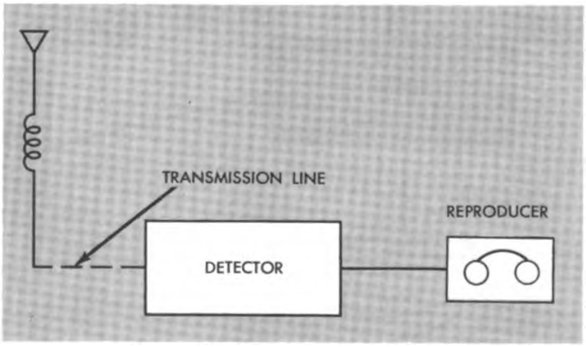
ANTENNA SYSTEM

The modulated carrier wave excites the transmitting antenna to produce electromagnetic radiations which correspond to the amplitude variations of the carrier wave. A transmission line (shown dotted in the illustration) carries the modulated radio-frequency energy to the transmitting antenna.

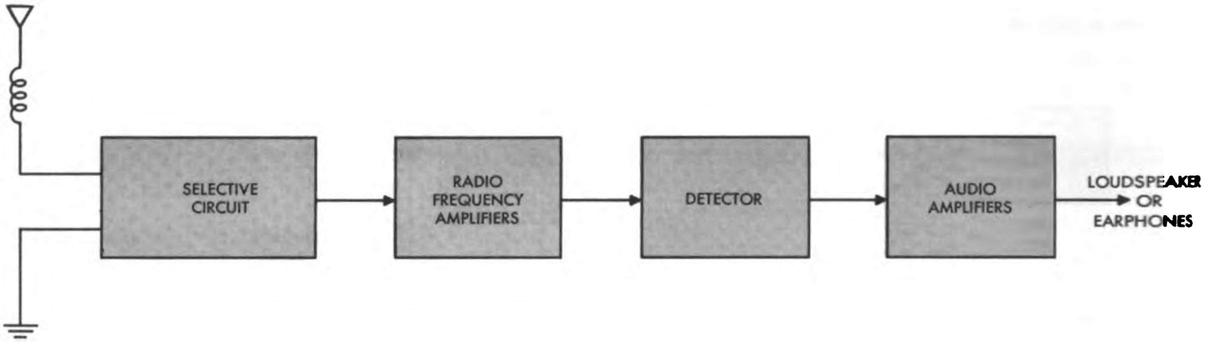
An important part of the antenna system is at the receiving end. This consists of a receiving antenna which picks up some of the electromagnetic radiations radiated from the transmitting antenna. A transmission line carries the received modulated RF carrier wave to the receiver where the information is removed from the carrier.

RECEIVERS

A receiver consists of the antenna and transmission line which collects the modulated carrier, a *detector* which removes the information, and a *reproducer* which reproduces the message in a form which can be seen or heard.



Block Diagram of Components of a Receiver



Block Diagram of a Radio Receiver

Radio Receiver

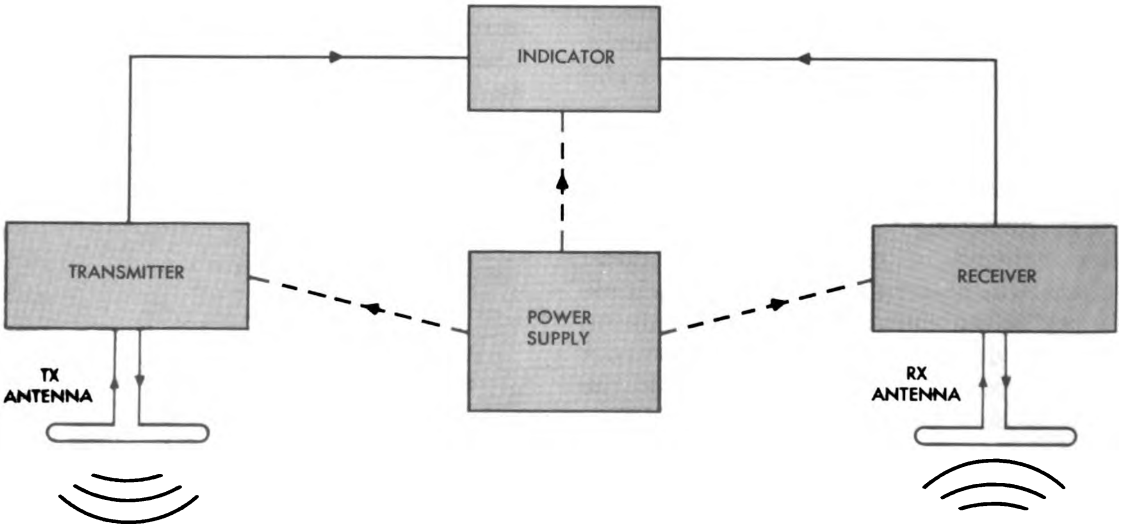
The function of the antenna and transmission line have already been described. The function of the detector is to detect or demodulate the modulated RF signal. This consists of separating the information signal from the carrier. If the detected signal is an audio signal, it is reproduced by earphones or a speaker.

A selective circuit is usually added to prevent the radio receiver from responding to a number of transmissions at the same time and interfering with the desired signal. You can see this in the block diagram above. Radio-frequency amplifiers are sometimes used to amplify the input RF signal before detection. Audio-frequency amplifiers may be used to amplify the signal after detection so that it can be reproduced by a loudspeaker or earphones.

Radar Receivers

The radar receiver and the radio receiver are essentially the same. While the radio receiver has as a reproducer a loudspeaker or headphones, the radar receiver has an indicator which displays signals visually. A cathode-ray tube is used as an indicator to display signals reflected from targets.

While the output of the detector of a radio receiver is an audio signal with which the carrier has been modulated, a detector of a radar



Block Diagram of SCR-718 Altimeter

receiver has as its output a video signal. The received radar signal results from the portion of the transmitted electromagnetic energy reflected from targets. Radio receivers are located at a distance from the transmitter sending the information. The modulated carrier therefore travels directly from the transmitter to the receiver. Radar receivers, on the other hand, are usually located near the transmitter. Therefore, the received radar signal travels from the transmitter to the target and is then reflected back to the receiver.

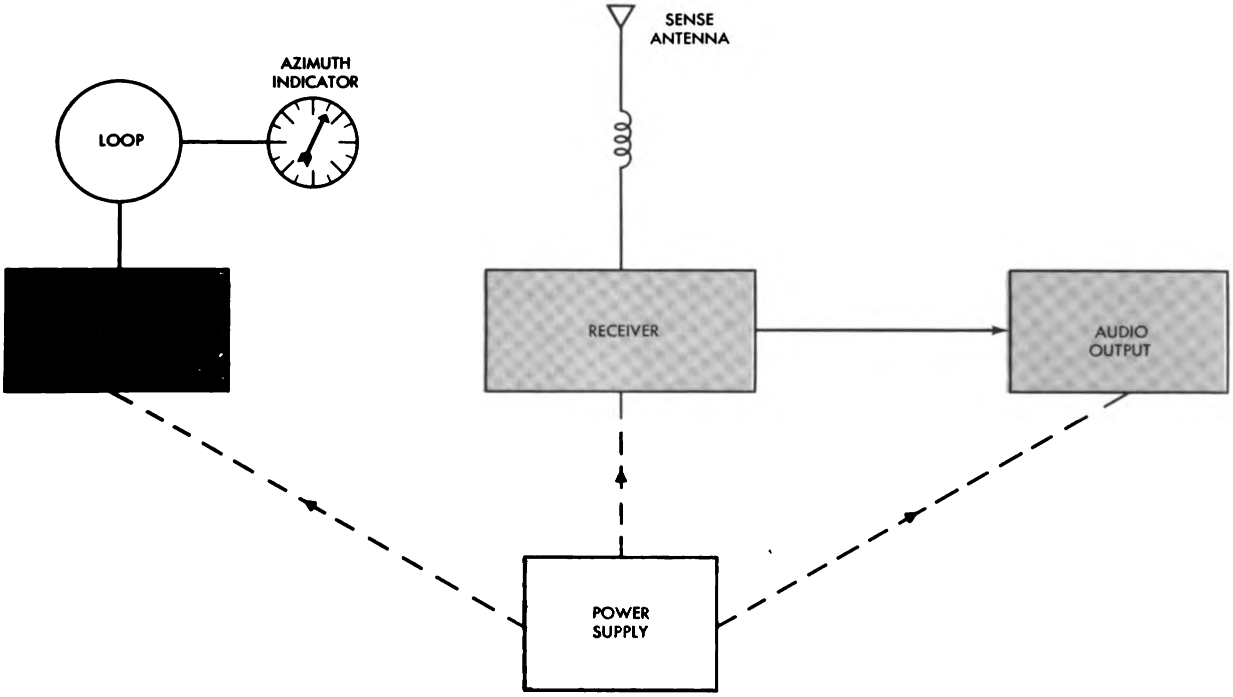
Typical Receivers

Some typical receivers with which you may be familiar are the SCR-718 Altimeter, the Radio Compass, and the AN/APN-9 Loran receiver. These are discussed briefly in the following sections.

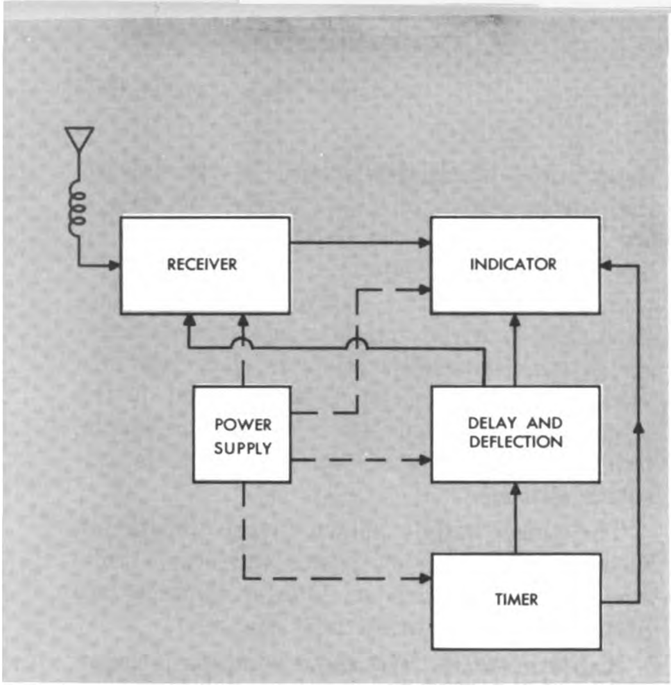
SCR-718 ALTIMETER. The altimeter is a complete system, consisting of a transmitter, a receiver, and an indicator as shown on page 194. The transmitter generates a radio-frequency signal which is then amplified and pulse-modulated. This means that the radio frequency is transmitted in short pulses at definite intervals. The transmitting antenna radiates these pulses of electromagnetic energy toward the earth. Some of these pulses of energy are reflected from the earth and picked up by the receiving antenna. The signals are then carried to the receiver by the transmission lines. The receiver amplifies and detects these RF signals; the result is a video signal. After this video signal is amplified, it is displayed on the screen of the indicator. From this display on the indicator, the operator can determine his absolute altitude.

The power supply shown in the illustration furnishes the necessary power and proper voltages for all components. Power supplies are discussed in detail in the next chapter.

RADIO COMPASS. The radio compass is essentially a receiver with its antenna system as shown below. The loop control rotates the loop antenna either manually or automatically toward the station to be received. The receiver amplifies and detects the signal picked up by the antenna. The audio output of the receiver is amplified and applied to headphones. A tuning device is used to tune the receiver to the desired radio frequency. The power supply furnishes the necessary power at the proper voltages.



Block Diagram of the Radio Compass



Block Diagram of an AN/APN-9 Receiver

AN/APN-9 LORAN. Loran is a system of navigation which uses synchronized, ground-based transmitting stations and receiver-indicators carried in aircraft, surface vessels, or land vehicles. The AN/APN-9 is an airborne Loran receiver-indicator used in Loran systems. As you can see in the block diagram above, pulsed signals transmitted by a pair of ground-based transmitting stations are picked up by the receiving antenna and carried to the receiver by transmission lines. The receiver amplifies, and detects the received signals and applies them to the cathode-ray tube for display. To make it possible to use the system for navigation, it is necessary to measure the time difference between the reception of the two signals. This is made possible by the timer, sweep delay, delay, and deflection circuits. The power supply provides power and necessary voltages for all components.

* * * * * *

This chapter has served merely to introduce to you the subject of transmission and reception in a general way. The block diagrams used in this chapter are intended to give you an over-all picture of what the purpose of each block is. The electronic circuits, contained within each block and which make possible the accomplishment of its function, are discussed in detail in the remainder of this manual. You will find a more complete discussion of transmission and reception in chapter 28 and chapter 29. Power supplies, principles of amplification, and principles of oscillation are discussed in the intervening chapters.

Chapter 25

Power Supplies

Vacuum tubes require various values of voltages for their different elements (such as filament, screen grid, plate). Except for the filaments which can be heated with AC or DC, the vacuum tube elements require DC voltages. A typical power supply for an electronic device is shown in the block diagram on page 204.

The *AC source* can be any source of alternating current. 115-volt, 60-cycle AC is typical of that encountered in homes. In aircraft applications, 400-cycle power sources are most common.

The *transformer* increases or decreases the AC voltage value of the voltage source. Step-up or step-down power transformers are used for this purpose. Such transformers are described on pages 111 and 112.

The *rectifier* is a diode (or diodes) which changes AC to pulsating DC. For lower values of voltage, selenium or copper-oxide rectifiers are frequently used.

The purpose of the *filter* is to change the pulsating DC obtained from the rectifier to a constant, smooth DC.

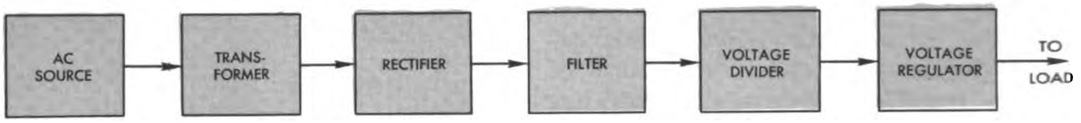
The *voltage divider* provides for several values of voltage from a single power supply. Voltage dividers already have been discussed on page 38.

The purpose of the *voltage regulator* is to keep the output voltage constant under changing load conditions.

The six elements of the power-supply system shown, are not all present in every application. The only essential parts are the AC source and the rectifier; the other elements contribute to the ability of the power supply to provide a wider range of output voltages which are smoother and more dependable. Since the other elements have already been discussed, this chapter presents a discussion of rectifiers, filters, and voltage regulators.

RECTIFIERS

In chapter 19, you learned that a diode vacuum tube acts as a rectifier. There are various types of circuits that are used for rectification — among these are half-wave rectifiers, full-wave rectifiers, bridge rectifiers, and voltage doublers.



Block Diagram of a Power Supply

Half-Wave Rectifier

The half-wave rectifier shown below is a circuit which employs a single diode to rectify an alternating current. T_1 is the power transformer whose primary is connected to an AC source. The rectified current flows through the load resistor.

The transformer secondary applies an AC voltage between the plate and cathode of the diode. However, current flows only during that half-cycle when the plate is positive with respect to the cathode. Current flows in only one direction as indicated by the arrows. However, the current is a pulsating type as shown by the waveshape. The frequency of the pulses is called *ripple frequency*. For each cycle of AC input there is one pulse or one ripple. Therefore, in this type of rectification, the ripple frequency is equal to the AC input frequency.

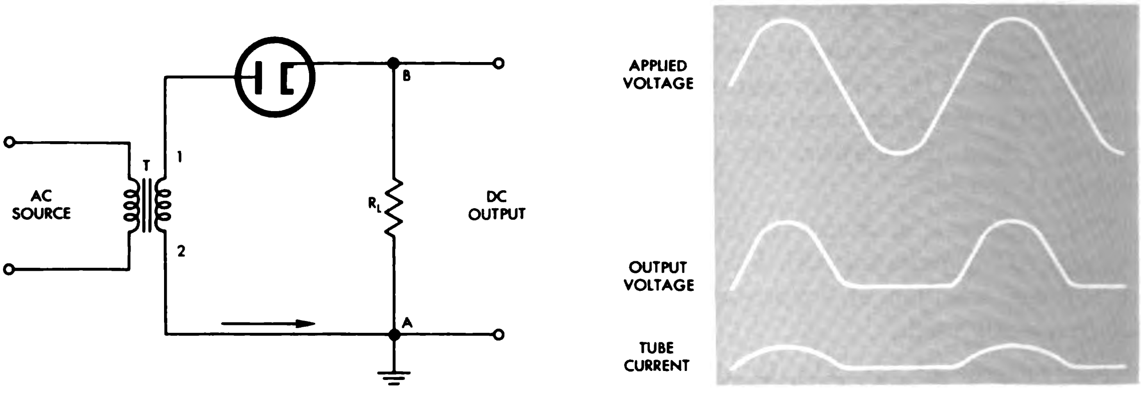
The current flows through R_L and causes a voltage drop to exist across it. Since the current flows from point A to point B in R_L, point A is negative with respect to point B. Because the current through R_L is pulsating, the voltage developed across it is also pulsating. From the waveshapes of the output voltage and current, you can see why this type of rectifier is called a half-wave rectifier.

If the tube is reversed so that the plate is connected to point B and the cathode to point 1 of the transformer secondary, current flows through R_L in the opposite direction (from B to A). Since A is at ground potential, the output voltage across R_L is negative in this case.

The output voltage of a half-wave rectifier drops to a low value if a load requiring much current is connected to the output terminals. Half-wave rectifiers are therefore used where current drain is low and high voltages are required. The high-voltage power supply circuit for an oscilloscope is usually a half-wave rectifier.

Full-Wave Rectifiers

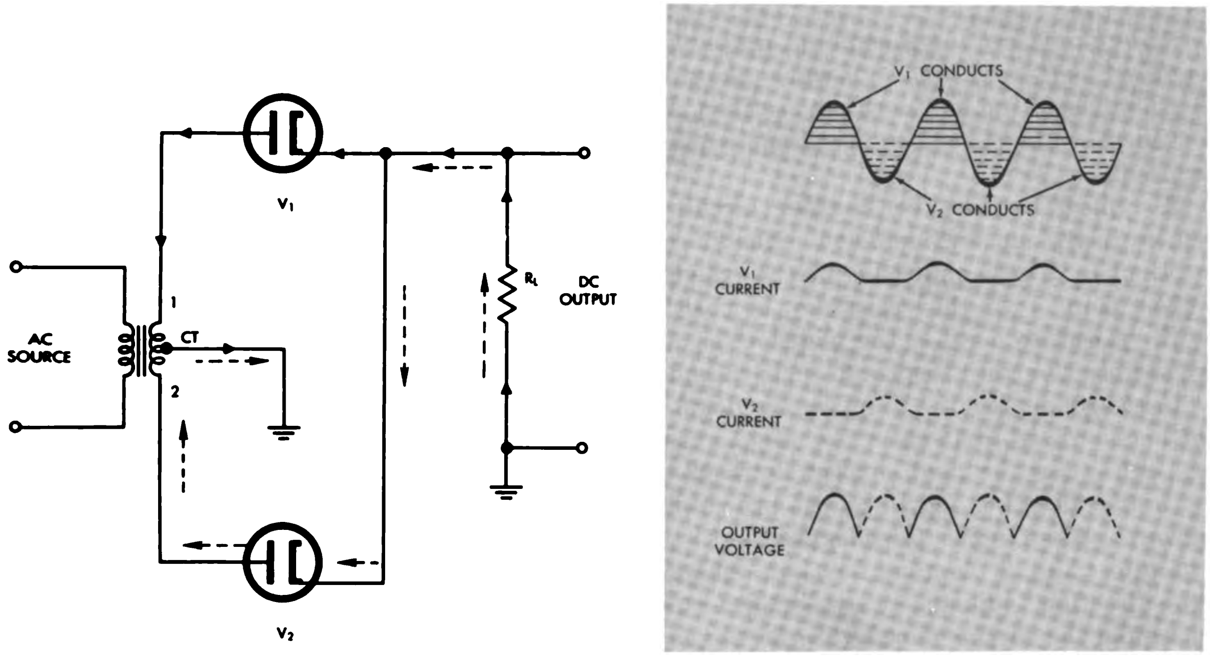
The half-wave rectifier finds many applications but it has several disadvantages. Some of these disadvantages are overcome by the use of a full-wave rectifier. A full-wave rectifier uses both halves of the AC cycle in its rectification process. A typical full-wave rectifier is shown on page 199. If you compare this circuit with that of a half-wave rectifier you will note two differences:



CIRCUIT

WAVESHAPES

Half-wave Rectifier



Full-wave Rectifier

1. The full-wave rectifier employs two diodes.
2. The high-voltage transformer secondary of the full-wave rectifier is center tapped (CT).

Note that the center tap is returned to ground and then through R_L to the cathodes of V_1 and V_2. Point 1 of the high-voltage winding is connected to the plate of V_1 and point 2 to the plate of V_2. Thus the AC voltage developed from 1 to CT is applied across V_1 and the voltage from 2 to CT is applied across V_2.

The two diodes conduct alternately since, at any given instant, one plate is positive and the other negative; a half-cycle later the polarity of the voltages is reversed. The direction and path of current through V_1 is shown by the solid arrows and through V_2 by the dotted arrows. Both currents flow through the load resistor in the same direction. The output and input voltage waveshapes are shown in the illustration. Note that the frequency of the pulses is twice the frequency of the AC source. If 60-cycle AC is rectified by a full-wave rectifier, the ripple frequency is 120 cps.

In most applications, the two diodes are included in one envelope. When V_1 is conducting, its plate is positive. The cathode of V_2 is at the same potential; however, its plate is negative at the same instant. Each tube must therefore be capable of withstanding twice the peak voltage applied to the tube in the reverse direction. This is known as the *inverse peak voltage* and must be taken into consideration in the design of a full-wave rectifier.

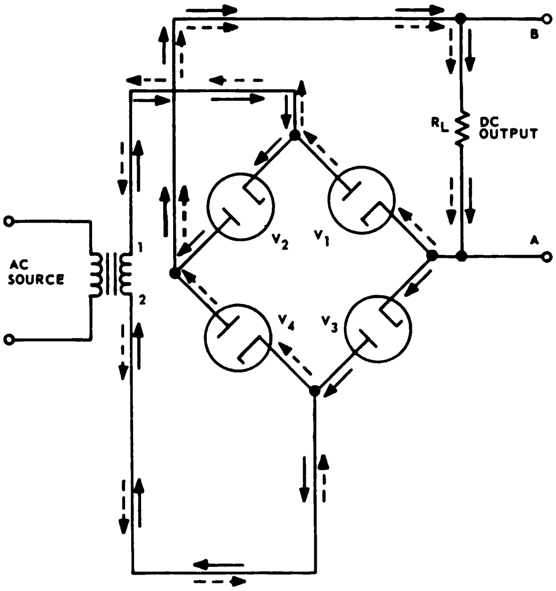
In full-wave rectifiers, where *half* of the high-voltage secondary is connected across each diode, the DC output voltage is proportional to *half* of the secondary's AC voltage. On the other hand, in half-wave rectifiers, the *entire voltage* of the high-voltage secondary is applied across the diode. The high-voltage secondary winding of a full-wave rectifier must therefore provide approximately twice the AC voltage to give the same DC output. A major disadvantage of full-wave rectifiers is that transformers of higher voltage output are necessary. At higher values of voltage, a difficult problem is created because of the excessive amount of insulation required. As a result, where thousands of volts are supplied, half-wave rectifiers are usually used. However, full-wave rectifiers are used in power supplies which require a high current drain at a lower voltage.

Bridge Rectifiers

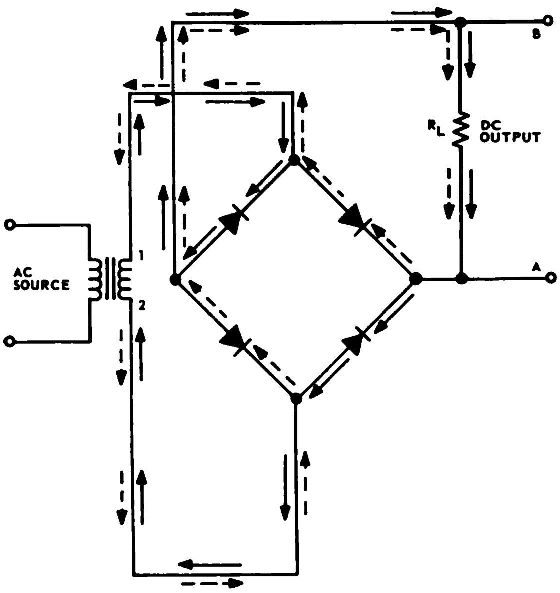
Bridge-type rectifiers combine the advantages of half-wave and full-wave rectifiers. Both halves of the AC wave are rectified, and the entire voltage of the high-voltage winding is used.

A typical bridge-type rectifier circuit is shown in the illustration below. Heater connections have been omitted from the drawing. When point 1 of the secondary is positive with respect to point 2, the plates of V_1 and V_4 are positive with respect to their cathodes, whereas the plates of V_2 and V_3 are negative with respect to their cathodes. As a result, V_1 and V_4 conduct and current flows in the path and direction shown by the broken-line arrows. When point 1 is negative with respect to point 2, current flows in the path and direction indicated by the solid arrows. Note that the current through R_L is always in the same direction, making point B negative with respect to point A. The ripple frequency of the output voltage of a bridge-type rectifier is twice the frequency of the input AC.

Copper oxide and *selenium* rectifiers are the principal types of metallic rectifiers. A copper oxide rectifier consists of a thin film of copper oxide on a copper plate. A selenium rectifier consists of a prepared film of selenium on a metallic substance such as iron. Both act as diodes in that they permit current to flow more readily in one direction than in the other. In each case, electrons flow from the metal to the film. Metallic oxide rectifiers cannot withstand high inverse voltages since the film is very thin. The breakdown voltage is increased, however, by building a stack of several such rectifiers in series. Another method of reducing the danger of breakdown is to arrange such rectifiers in a bridge circuit as shown in the illustration above. The two current paths are shown with solid and broken-line arrows as before.



Bridge Rectifier Using Diodes



Selenium Bridge Rectifier

Voltage Doublers

In some smaller radio receivers, higher DC voltages are sometimes obtained without the use of a transformer. The circuit that accomplishes this is called a *voltage-doubler* rectifier. As its name implies, the voltage doubler rectifies and approximately doubles the AC input voltage. In order for a voltage-doubler rectifier to function properly, the amount of current drawn from it must be kept small. In radar circuits, a voltage doubler is often used as a high-voltage source where little current is required. The circuit for a basic voltage doubler is shown to the right. Two separate diodes are shown; a twin diode may be used if each diode has a separate cathode. Twin diodes, like the 5U4, cannot be used, because there is only one cathode for both plates. The 25Z5 is an example of a tube that may be used as a voltage doubler.

When point 1 of the transformer in the illustration is positive with respect to point 2, V_1 conducts, charging the capacitor C_1. The path in which current flows is indicated by dotted lines on the figure. When C_1 is charged to the maximum applied voltage, no further current flows in V_1, since the voltage across C_1 makes the plate of V_1 negative with respect to its cathode. If the voltage across C_1 is equal to the maximum reached by the high-voltage secondary, the plate *cannot be positive* with respect to the cathode, and the tube cannot conduct. Similarly, when point 1 is negative with respect to point 2, V_2 conducts until C_2 is charged to maximum applied voltage. Thus, the voltage across each capacitor is maximum, and the output voltage from A to B is twice the maximum applied voltage. A bleeder resistor is used to discharge the capacitors when the equipment is turned off. It is made large to prevent the capacitors from discharging too much between charging alternations.

When a load resistor is connected across the output terminals, the capacitors discharge through it. Thus, the capacitors lose some of their charge, and in so doing, the voltage across them decreases until the diodes can recharge them. The amount of the decrease in voltage depends upon the amount of current drawn by the load resistor. The output voltage becomes less as the capacitors discharge. The load must therefore be light if the output voltage is to be fairly constant. The relationship between the input and output voltages and currents in the voltage-doubler circuit are shown. Note that the ripple frequency is twice the frequency of the AC input. You can consider a voltage doubler to be two half-wave rectifiers connected so that their outputs are added in series.

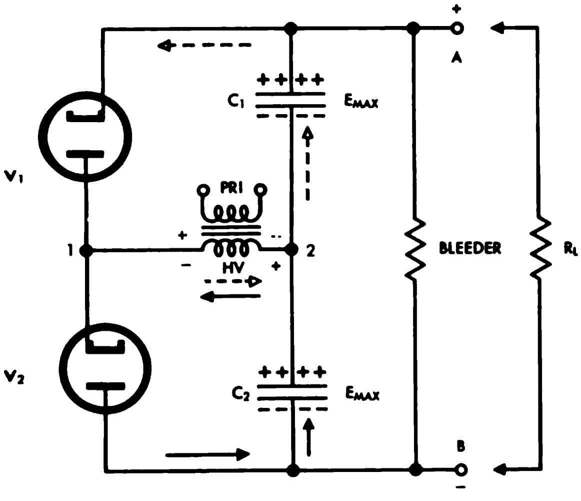
FILTERS

You have seen that the rectifier changes AC voltage to DC voltage. However, the DC voltage output of the rectifier is a pulsating DC voltage; that is, it has pulses or ripples. These ripples may be considered to be an AC voltage superimposed on a DC voltage. If the output from a rectifier is to be used to provide voltages for certain vacuum tube circuits, the ripple must be removed or reduced, otherwise distortion may occur. *A filter* of a power supply is a device which removes or reduces the magnitude of the AC component (ripple) of a rectifier so that only the DC component is effective.

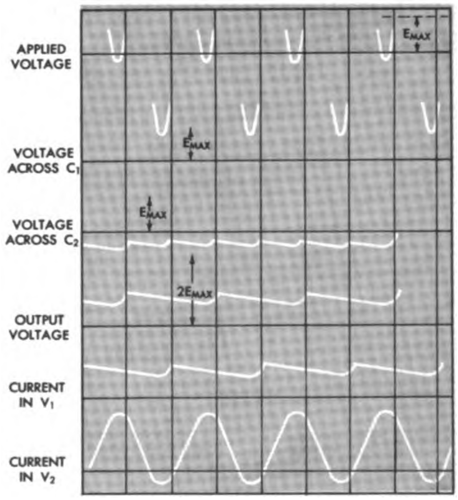
Individual Reactances as Filters

A reactance which opposes a change in voltage (or current) by storing energy and then releasing this energy back to the circuit may be used as a filter.

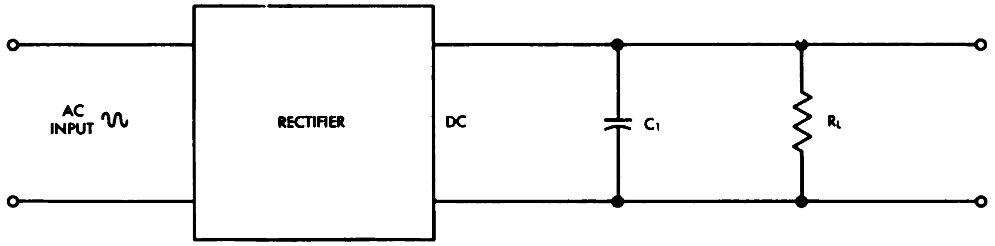
EFFECT OF CAPACITANCE. You have seen that a capacitance opposes a voltage change across its terminal by storing energy in its electrostatic field. Whenever the voltage tends to rise, the



A Voltage Doubler



Wave Forms in a Voltage Doubler



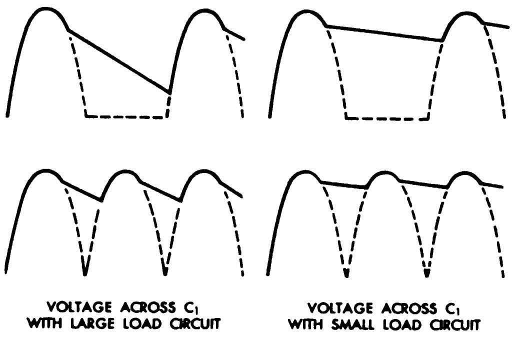
A Capacitor as a Filter

capacitor converts this voltage change to stored energy. When the voltage tends to fall, the capacitor converts this stored energy back to voltage. The use of a capacitor for filtering the output of a rectifier may be seen in the illustration above. The rectifier is shown as a block in the illustration, and the capacitor C_1 is connected in parallel with the load R_L.

The capacitor C_1 is chosen to offer very low impedance to the AC ripple frequency and very high impedance to the DC component. The ripple voltage is therefore bypassed to ground through the low-impedance path, while the DC voltage is applied unchanged to the load. The effect of the capacitor on the output of the rectifier may be seen in the waveshapes below. Dotted lines show the rectifier output; solid lines show the effect of the capacitor. Half-wave and full-wave rectifier outputs are shown. The capacitor C_1 charges when the rectifier voltage output tends to increase and discharges when the voltage output tends to decrease. In this manner, the voltage across the load R_L is kept fairly constant.

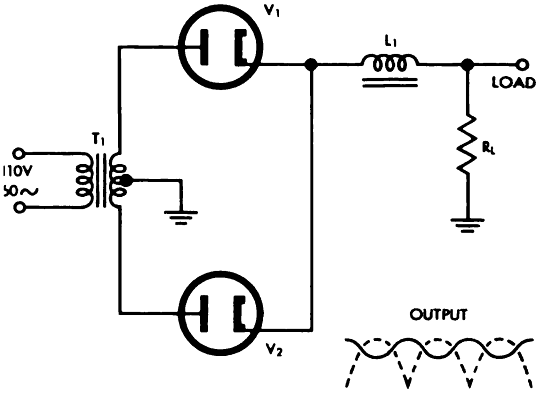
EFFECT OF INDUCTANCE. An inductance may be used as a filter, because it opposes a change in current through it by storing energy in its electromagnetic field whenever current tends to increase. When the current through the inductor tends to decrease, the inductor supplies the energy to maintain the flow of current. The use of an inductor for filtering the output of a rectifier may be seen in the illustration below. Note that the inductor L_1 is in series with the load R_L.

The inductance L_1 is chosen to offer high impedance to the AC ripple voltage and low impedance to the DC component. Therefore, for the AC ripple, a very large voltage drop occurs across the inductor and a very small voltage drop across the load R_L. For the DC component, however, a very small voltage drop occurs across the inductor and a very large voltage drop across the load. You can see the effect of an inductor on the output of a full-wave rectifier in the output waveshape. Note how the ripple has been attenuated in the output voltage.

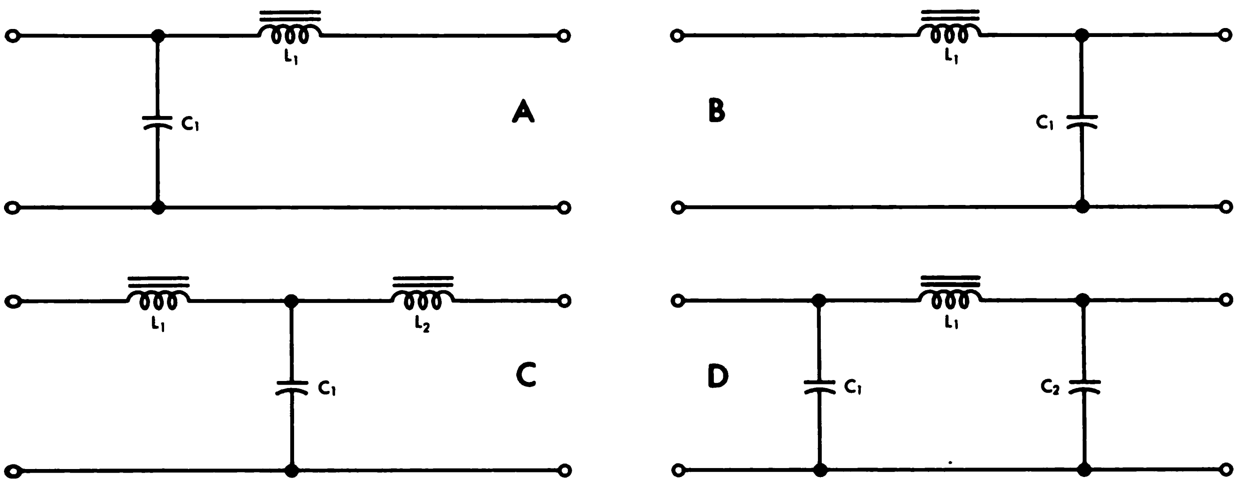

VOLTAGE ACROSS C_1 WITH LARGE LOAD CIRCUIT

VOLTAGE ACROSS C_1 WITH SMALL LOAD CIRCUIT

Effect of Capacitor on Rectifier Output
Half-wave and Full-wave Outlets Shown



An Inductor as a Filter



LC Filter Sections

LC Filters

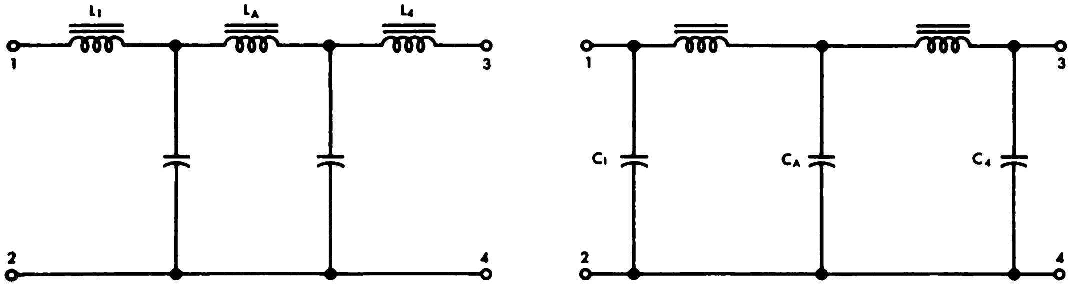
Capacitors and inductors are combined in various ways to provide more satisfactory filtering than can be obtained with a single capacitor or inductor. Several combinations are shown schematically. Note that the L-, or inverted L-, type and the T-type filter sections resemble schematically the corresponding letters of the alphabet. The pi-type filter section resembles the Greek letter pi (π) schematically.

Each of the filter sections shown above are similar in that the inductances are in series and the capacitances in parallel with the load. The inductances must therefore offer a very high impedance and the capacitors a very low impedance to the ripple frequency. Since the ripple frequency is comparatively low, the inductances are iron-core coils having large values of inductance (several henries). Because they offer such high impedance to the ripple frequency, these coils are called *chokes*. The capacitors must also be large (several microfarads) to offer very little opposition to the ripple frequency. Because the voltage across the capacitor is DC, electrolytic capacitors are frequently used as *filter capacitors*. Be sure you observe the correct polarity in connecting electrolytic capacitors.

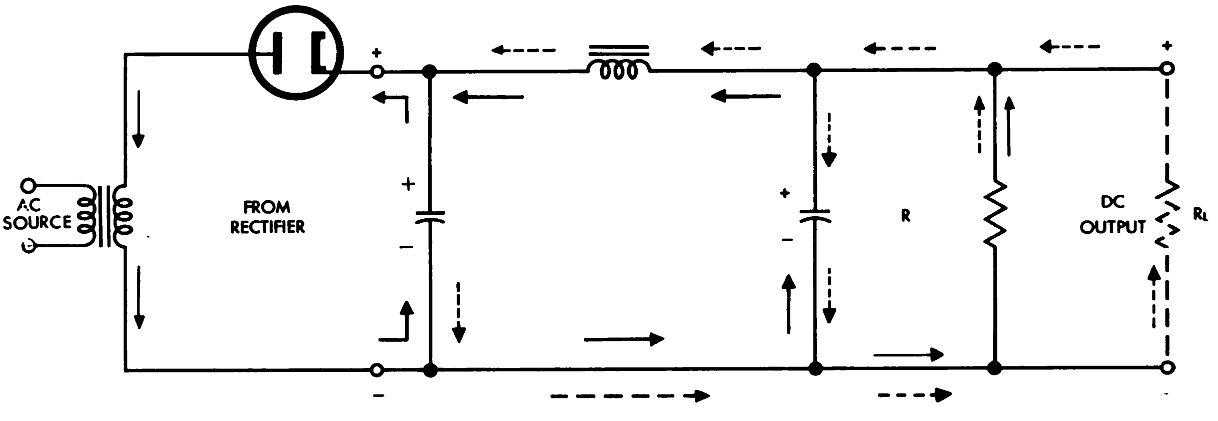
More than one section of a given type of filter may be combined to improve the filtering action. Some two-section filters are shown in the illustration below.

LC filters are also classified according to the position of the capacitor and inductor. A *capacitor-input* filter is one in which the capacitor is connected directly across the output terminals of the rectifier. Filter sections A and D are examples of capacitor-input filters. A *choke-input* filter is one in which a choke precedes the filter capacitor. Filter sections B and C are examples of choke-input filters.

CAPACITOR-INPUT FILTERS. The most common type of filter used with half-wave rectifiers is the capacitor-input, pi-type filter shown in the



Filters of Two Sections

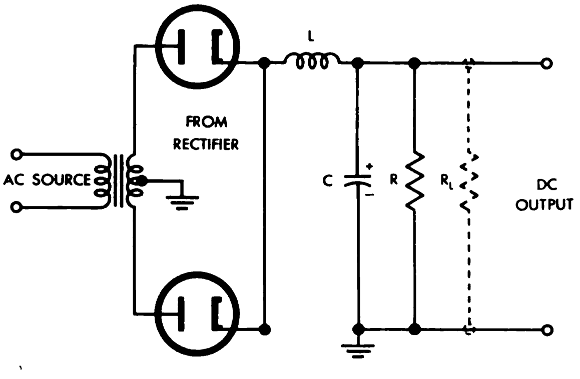


Half-wave Rectifier with Capacitor-input Filter

illustration above. When the electrons flow in the path shown by the solid arrows, the capacitors are charged as shown. If R_L is not connected and the capacitors are assumed to have no leakage, the output voltage is equal to the peak value of the AC voltage applied from the secondary of the transformer. If R_L is connected, the capacitors discharge through it, as shown by the dotted arrows. The extent to which the capacitors discharge depends upon the value of the load. The voltage falls rapidly as the current drain increases.

The resistor R shown as part of the filter is called a *bleeder resistor*. It is used to discharge the capacitors when the equipment is turned off. The resistor should be large; it should not draw more than 10% of the rated current.

The output voltage of a rectifier with a capacitance-input filter is approximately peak value if the current drain is low, but the load must be extremely light. Because the output voltage depends on the value of the load, a capacitance-input filter has poor voltage regulation. If a full-wave rectifier is used with a capacitance-input filter, there is a slight improvement in the filtering action. However, the voltage regulation is still poor.



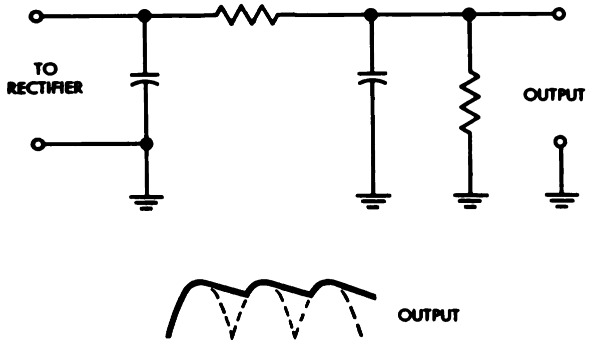
Full-wave Rectifier with Choke-input Filter

CHOKE-INPUT FILTERS. Choke-input filters provide better voltage regulation than capacitance-input filters. As a result, choke-input filters are used to a great extent in power supplies for transmitters and other electronic circuits which require considerable power. However, the choke-input filter requires full-wave rectification and gives less DC output voltage for a given AC input.

A full-wave rectifier with a choke-input filter is shown in the illustration. The pulsating current from the full-wave rectifier flows through the load and the choke. The choke opposes the changes in current and reduces the magnitude of the ripple. The capacitor filters the AC component further. When a small amount of current is drawn, the voltage drops rapidly and then remains fairly constant over a wide range of load. This type of filter gives good regulation under changing load conditions or heavy drain.

RC Filters

In many applications, a resistor is used to replace the coil in a filter to save expense. A disadvantage of this type of filter is that the resistor offers the same impedance to the DC



RC Pi-type Filter

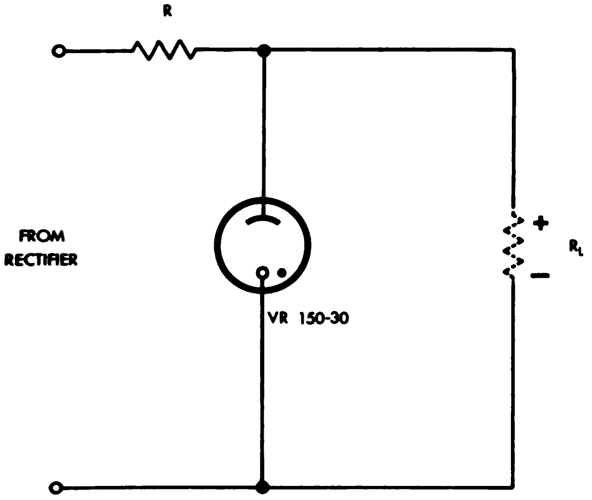
voltage as to the AC component. In addition, current flow through the resistor causes power to be dissipated in the form of heat.

VOLTAGE REGULATORS

Voltage regulators are circuits designed to maintain the output voltage of a power source at a constant, predetermined level, even though the input voltage or the load varies.

VR Tubes

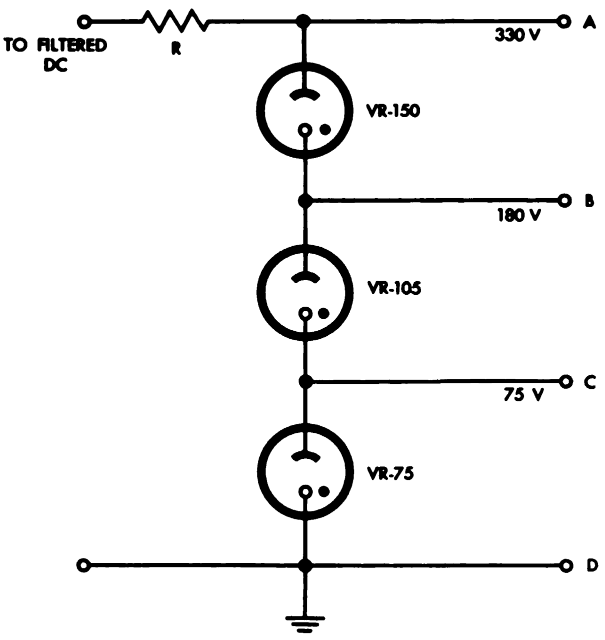
As you recall from your study of gas tubes, a VR tube has a constant voltage drop across it if the current is held within limits. In the circuit shown below, the output voltage is regulated by the VR 150-30 tube at 150 volts as long as the current through the tube does not exceed 30 ma. A minimum current of 5 ma. must be maintained to keep the tube from extinguishing. In other words, the current through the tube must be maintained between 5 ma. and 30 ma. to keep the output voltage regulated.



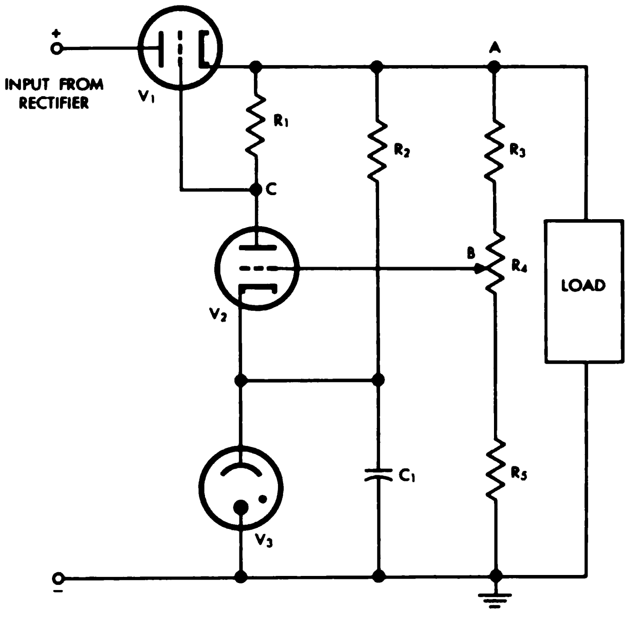
Gas Tube Voltage Regulator

By proper selection of limiting resistor R, the output voltage is maintained at 150 volts regardless of changes in input voltage from the rectifier or of changes in load. If the input voltage rises, more current flows through the VR tube and a greater voltage drop occurs across R. This greater voltage drop keeps the voltage across the load R_L constant. A decrease in input voltage causes a smaller voltage drop across R and so keeps the output voltage constant. A change in load is compensated for by a change in current flowing through the VR tube. This, in turn, keeps the voltage drop across R constant and therefore maintains a constant output voltage.

One of the major disadvantages of VR tubes is their comparatively low voltage rating. This disadvantage may be overcome by connecting several VR tubes in series as shown in the illustration below. This circuit provides for three regulated voltages: 330 volts from A to ground,



VR Tubes in Series



Electronic Voltage Regulator

R_1	Plate load resistor for V_2. Plate current of V_2 flows through it.
R_2	Limiting resistor for V_3. Limits the current flow through V_3.
R_3 R_4 R_5	Series bleeder resistors. R_3 and R_5 are large to maintain the current through the potentiometer R_4 at a low value.
R_4	Potentiometer which adjusts the grid voltage of V_2 and sets the output voltage of the power supply.
C_1	Large filter capacitor. Stabilize the voltage across the VR tube.
V_3	Voltage regulator tube. Maintains the cathode potential of V_2 constant.
V_2	A DC amplifier. It adjusts the bias of V_1 as the output voltage tends to change.
V_1	Regulator tube. The DC resistance of this tube changes to maintain the output voltage constant.

Summary of Voltage Regulator Components

180 volts from B to ground, and 75 volts from C to ground. Note that a limiting resistor R is used also in this circuit to keep the current flowing through the tubes between 5 ma. and 30 ma.

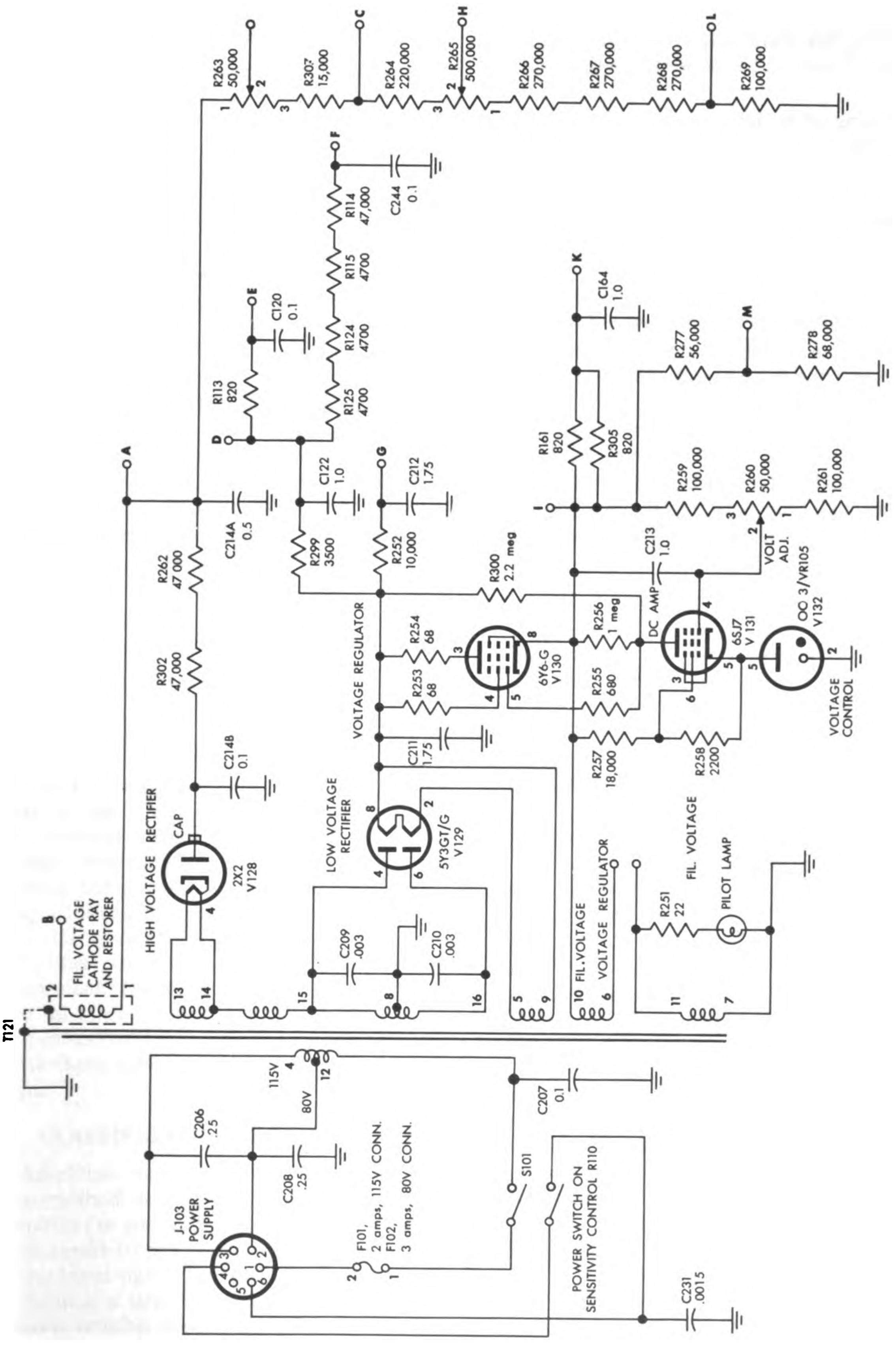
Electronic Voltage Regulators

Electronic voltage regulators operate on the principle that a vacuum tube functions as a variable resistor. When a tube conducts, its DC plate resistance is the plate-to-cathode voltage divided by the DC current flowing through the tube. An increase in bias decreases the current flow and increases the DC plate resistance of a tube; a descrease in bias increases the current flow and decreases the DC plate resistance of a tube.

The electronic voltage regulator shown to the left is capable of very close regulation at a level which may be set by varying the potentiometer setting. This circuit contains a vacuum tube V_1 in series with the load. The voltage across the load is regulated by controlling the conduction of V_1. Thus, V_1 acts as a variable resistor that automatically adjusts itself to the correct value. V_3 is a VR tube, which maintains the cathode of V_2 at a fixed, positive potential. The voltage-divider system, composed of R_3, R_4, and R_5, is arranged so that the variable arm of R_4 can be adjusted to a positive voltage sufficiently low to bias V_2 so that it operates on the linear portion of the E_g-I_p curve. R_1 is the plate load resistor for V_2, which is connected in series with VR tube V_3. The purpose of this resistor is to absorb any changes of voltage and keep the cathode of V_2 at a fixed potential.

If the output voltage (point A) tends to rise due to an increase in input voltage or to a decrease in current drawn by the load, the voltage at point B also tends to rise. The voltage at B is the voltage applied to the grid of V_2 and therefore determines how much V_2 will conduct. If V_2 conducts more, there is an increase in current through R_1 and a decrease in voltage at point C. This decrease in voltage is applied to the grid of V_1 and causes it to conduct less. This increase in the resistance of V_1 counteracts the tendency of the rise in output voltage and keeps the voltage constant.

If the output voltage should tend to fall, the voltage at B tends to drop, causing a rise in



Power Supply for AN/APN-9 Receiver

voltage at C. This increase in voltage at C causes V_1 to conduct more and increases the output voltage.

The setting of B determines the bias of V_2 and therefore the current flowing through it. As this current flows through R_1 it determines the bias on V_1. The DC plate resistance of V_1 determines the output voltage. Hence, the output voltage is adjustable, within limits, by setting the movable tap, point B, of potentiometer R_4. Similar circuits are used in the power supplies of radar equipment. Because of the great amount of current flow required in some sets, you may find two or more tubes in parallel to serve the purpose of V_1. The rectifier and filter circuit must be designed for greater current capacity than the load is expected to draw since the regulator also draws current. The table in the illustration on page 212 summarizes the function of each of the components of the voltage regulator circuit.

A TYPICAL POWER SUPPLY

The power supply shown on page 213 is typical of those found in many radar sets. It is the complete power supply for the Loran receiver used as an aid to navigation. Note that this power supply has all six elements of the block diagram shown on page 204.

The *AC source* is 115 volts or 80 volts obtained from the input jack. The *transformer* T 121 has a tapped primary and many secondaries for high voltage and filament voltages. There are two *rectifiers*. The high-voltage rectifier is a half-wave rectifier and the low-voltage rectifier is a full-wave rectifier. There are various *filters*. These are all Pi-type, RC filters. The full-wave rectifier has an electronic *voltage regulator* consisting of tubes V 130, 131, and 132. Various voltages can be obtained from the *voltage dividers* of both rectifiers.

Study this power supply and try to identify the various elements discussed.

Chapter 26

Principles of Amplification

You have seen in your study of vacuum tubes that a signal may be amplified, or increased, by electronic means. An *amplifier* is a device, usually containing a vacuum tube, which increases the strength of an electrical signal. It may consist of a single stage (one vacuum tube and its associated circuit) or several stages. Amplifiers are necessary in nearly all electronic equipment. Radio, radar, television, telephone, and all other types of electronic devices employ amplifiers to raise the strength of extremely weak input signals so that they can be heard or seen.

CLASSIFICATION OF AMPLIFIERS

Amplifiers may be classified in several ways. One method is to classify amplifiers as *voltage* amplifiers or *power* amplifiers. A voltage amplifier is designed to produce a large voltage increase in the input signal. A power amplifier is designed to furnish a large amount of power to a loudspeaker or other power-consuming device. Power amplifier stages have a relatively large current flow. If voltage amplification occurs in a power amplifier, it is merely a secondary consideration. When greater amplification than can be obtained from a single tube is desired, several stages are combined to form an amplifier. The output of each stage is fed to the input of the succeeding stage until sufficient amplification is achieved. A complete amplifier usually consists of several voltage amplifier stages followed by a power amplifier stage.

Another method of classifying amplifiers is *according to the amount of bias voltage* applied to the vacuum tube. There are three general bias classifications: *class A*, *class B*, and *class C*. A *class A* amplifier is one in which the grid bias and signal voltages are such that plate current flows at all times. A *class B* amplifier is one in which the grid bias is approximately equal to the cutoff voltage of the tube. A *class C* amplifier is one which is biased appreciably below the cutoff voltage of the tube.

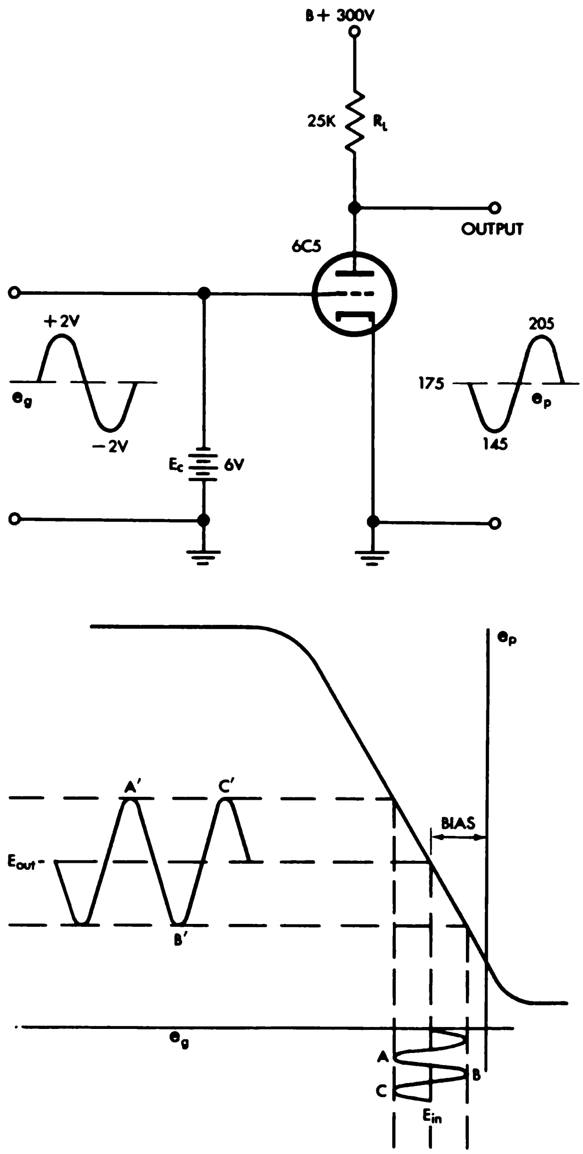
The *coupling between stages* may also be used to classify amplifiers. There are *direct coupled*, *RC coupled*, *impedance coupled*, and *transformer coupled* amplifiers. The direct coupled amplifier has the plate of one tube connected directly to the grid of the next. The RC coupled amplifier employs a resistor and a capacitor as a coupling circuit between the plate of one tube and the grid of the next. The impedance coupled amplifier normally employs an inductive plate load with an RC coupling circuit between the tubes. Transformers, when used as coupling between tubes, may be either tuned or untuned.

Amplifiers may also be classified *according to the band of frequencies* they are designed to amplify. There are four main divisions: *direct current* amplifiers, *audio frequency* amplifiers, *radio frequency* amplifiers, and *video frequency* amplifiers. Direct current amplifiers amplify slow or irregular changes in DC. Audio (frequency) amplifiers are designed to amplify frequencies from about 15 cycles per second to approximately 15,000 cycles per second. All amplifiers designed to handle frequencies above the audible range may be classed as radio frequency amplifiers. An RF amplifier, however, will not amplify the complete range of radio frequencies but only a small band of this complete range. Video (frequency) amplifiers are wide band amplifiers used in radar and television where the signals involved contain a wide range of frequencies. The various types of amplifiers are discussed more fully in the remainder of this chapter.

CLASS A AMPLIFIERS

A class A amplifier is one in which the grid bias and alternating signal voltage are chosen so that plate current flows throughout the entire cycle. Class A amplifiers are biased to operate on the linear portion of the dynamic curve. A typical class A amplifier is shown in the illustration.

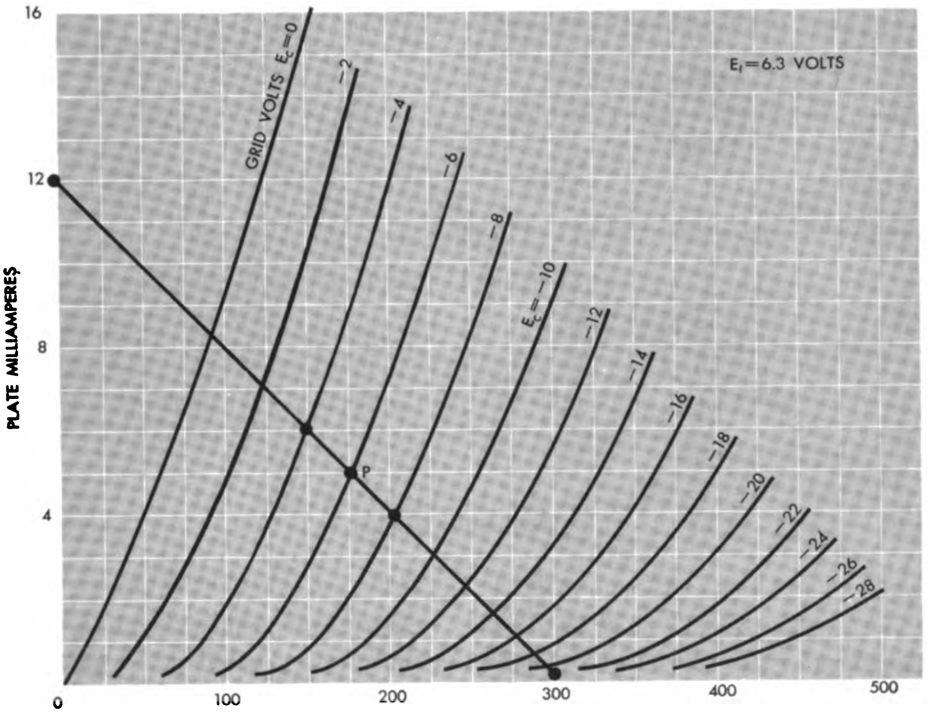
If a 4-volt, peak-to-peak signal is applied to the grid of the tube shown, an amplified output is obtained as plate voltage. The load line for this circuit is shown on page 217. The operating point of this tube, with no signal applied, is the intersection of the load line with the 6-volt bias line. As the signal goes through a cycle, grid voltage varies from −6 volts to −4 volts, back to −6 volts, down to −8 volts, and back to −6 volts. The output signal at the plate varies from 175 down to 145 volts and back up to 205 volts. Since the plate voltage varies 60 volts peak-to-peak while the grid voltage signal is 4 volts peak-to-peak, the stage has amplified the voltage of the input signal 15 times.



A Class A Amplifier

Note that the output signal varies the same below as above the operating point. It is, therefore, a faithful reproduction of the input signal. A class A amplifier introduces very little distortion. In addition, current flows throughout the cycle and without any input signal.

Class A amplifiers have low efficiency; that is, the ratio of power output to power input is low.



Load Line for Class A Amplifier Shown

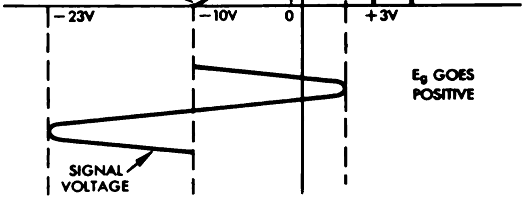
The efficiency of most practical class A amplifiers lies somewhere about 20%. Theoretically, a class A amplifier has a maximum efficiency of about 50%. Class A amplifiers are used to amplify audio and radio frequencies in radio, radar, and sound systems.

CLASS B AMPLIFIERS

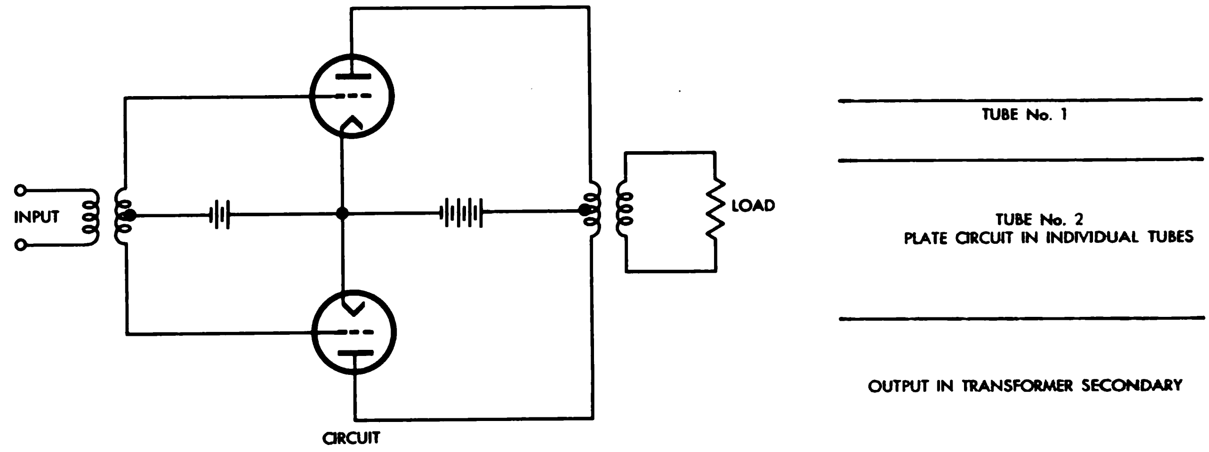
A class B amplifier is one in which the grid bias is close to cutoff. Plate current therefore flows only during the positive half-cycle of the applied signal voltage. The efficiency is high and the current consumption is less in a class B amplifier than in one operating class A. Power loss in class B amplifiers is low for two reasons. First, plate current does not flow when there is no signal applied to the grid; thus, there can be very little power wasted during the non-operating periods. Secondly, plate current flows only during the positive half of the input cycle. This means that the average current flow will be only 32% of the peak current in the stage.

The illustration to the right shows the relation between the grid voltage and plate current in a tube operated class B. Note that plate current flows only during the positive half of the signal voltage. Grid current flows only during the time when the grid is driven positive.

Class B amplifiers are used most as power amplifiers. As power amplifiers, their power output is proportional to the square of the grid-excitation voltage. Amplitude distortion in the output of a class B amplifier operating with the



Class B Amplification


INPUT
LOAD
CIRCUIT
TUBE No. 1
TUBE No. 2
PLATE CIRCUIT IN INDIVIDUAL TUBES
OUTPUT IN TRANSFORMER SECONDARY

Push-Pull Amplifier

proper load resistance depends upon the departure of the characteristic curve from straight lines and upon the operating point. Refer to the class B amplification curve shown. Obviously, the half-cycle output is far from the distortionless reproduction of the full cycle input. One-half of the cycle is missing. This missing cycle is supplied by one of two methods — by an additional class B amplifier, or by the flywheel action of a resonant circuit. Distortion of the positive peaks of grid voltage may occur if the grid goes positive. This distortion is caused by grid current flow.

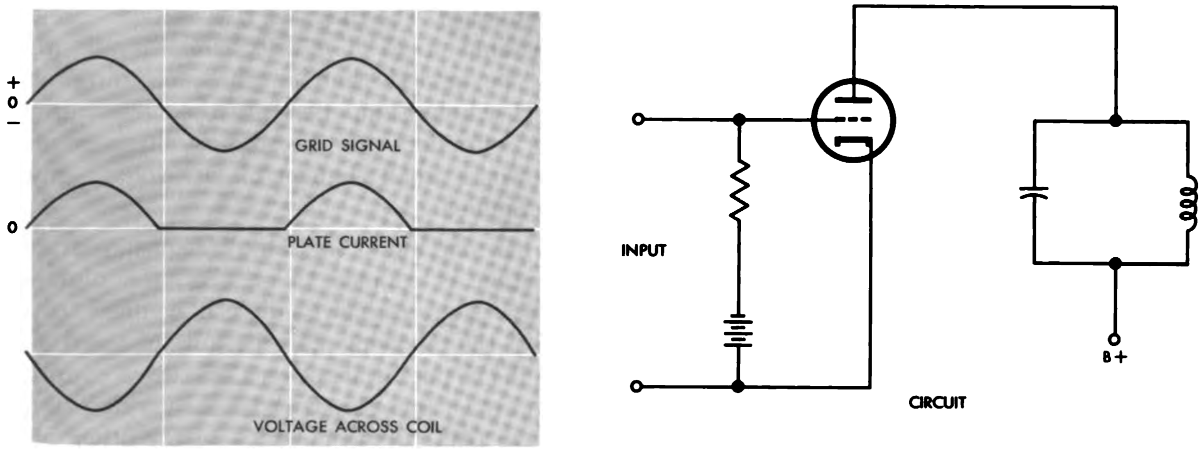
Push-Pull Class B Amplifier

The use of two tubes to supply both halves of the input cycle in the output constitutes a push-pull class B amplifier. One tube operates during the first half-cycle of the AC signal voltage and the other tube during the other half-cycle. Since plate current flows during one-half-cycle in one tube and during the next half-cycle in the other, the plate current wave forms can be combined in the load circuit. The load circuit of the push-pull amplifier is the center-tapped primary of the output transformer. During one cycle, one tube generates a voltage across the transformer winding. During the next half-cycle, the other tube generates a voltage of opposite polarity across the winding. Since the plate currents of the two tubes flow in opposite directions through their respective halves of the transformer primary winding, the voltages across the primary windings of each tube combine in the secondary to produce a reasonably undistorted output of the input AC signal voltage.

Single-Tube Class B Amplifier

A class B amplifier can be used successfully in an RF amplifier stage having a parallel-tuned circuit as the plate load. The parallel-tuned circuit is sometimes called a tank circuit, because it has the ability to store power. When it is used as the plate load of a single-ended, class B amplifier stage as shown on page 219, the capacitor in the parallel-tuned circuit is charged by the flow of plate current through the load on the positive half-cycles. Although no current flows through the tube on the negative half-cycles of the applied signal voltage, the capacitor discharges into the inductor during this period, and supplies the missing half-cycle in the output voltage. This so-called flywheel effect of the tank circuit occurs only when the resonant frequency of the parallel-tuned circuit is equal to the frequency of the applied signal voltage.

The maximum amount of useful power available from a single power tube is limited by the amplitude distortion caused by the introduction of undesired harmonic frequencies in the amplifier circuit. A harmonic frequency is a multiple of any given frequency. For example, the second



Class B Amplifier with Resonant Tank

harmonic of 1000 cps is 2000 cps, the third harmonic is 3000 cps, and so on. Second and higher harmonics are generated by a vacuum tube when its grid voltage-plate current changes are nonlinear.

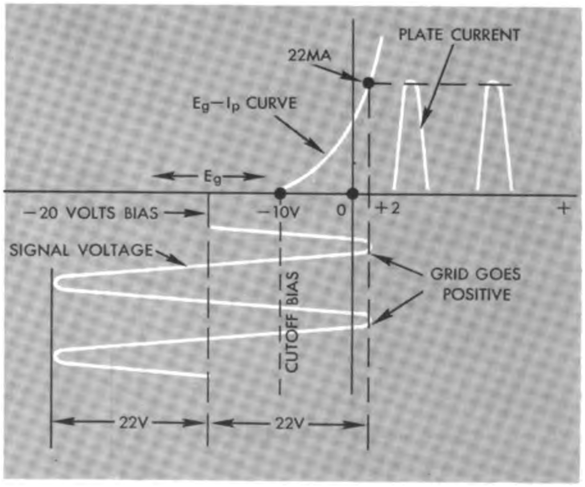
The theoretical maximum efficiency that is possible with a class B power amplifier is 78.5%. In most practical applications, however, the efficiency attained is between 60% and 65%. Class B amplifiers are used in transmitters as voltage amplifiers for amplitude-modulated waves.

CLASS C AMPLIFIER

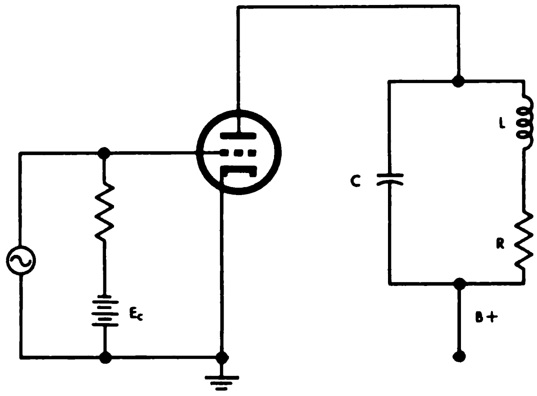
A class C amplifier is an amplifier in which the grid bias is appreciably greater than the cutoff value. When no alternating voltage is applied to the grid, the plate current is zero. When an alternating voltage is applied to the grid, plate current flows for appreciably less than one-half cycle as shown to the right.

Except for grid bias, class C amplifiers operate in the same way as class B amplifiers. Since current flows for only a small part of a cycle, the distortion in class C amplifiers is very great. The grid of a class C amplifier usually swings sufficiently positive to allow saturation-point current to flow through the tube. As a result, the plate output waves are not free from harmonics. Suitable means must therefore be provided to remove them from the output. One means of doing this is demonstrated by the illustration at the top of the next page. The tuned tank circuit shown offers a high impedance to the operating frequency with a power factor nearly equal to unity. The tank presents a low impedance to the harmonic frequencies, however, and causes them to be attenuated. The signal frequency is therefore amplified to a much greater extent than are the harmonic frequencies.

The efficiency of a class C amplifier is from 60% to 80%. The high efficiency of the class C amplifier is due to the fact that the instantaneous potential of the plate is low compared to the plate supply voltage when plate current flows.



Class C Amplification



Preventing Distortion with a Resonant Tank

Therefore, energy is supplied to the plate circuit only when the current causes a voltage drop across the tuned circuit.

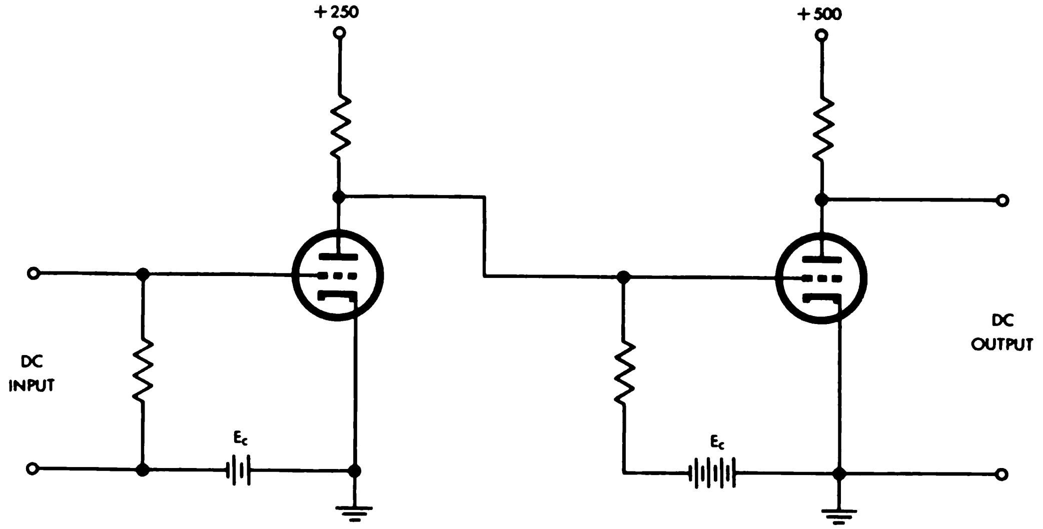
Class C amplifiers are used as RF amplifiers and modulators in transmitters and in high-frequency equipment where it is necessary to deliver appreciable power.

TYPES OF COUPLING

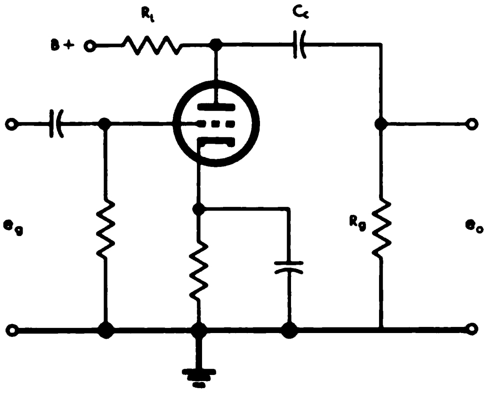
The type of coupling used between stages of an amplifier is often used as a method of classifying them. A *direct-coupled* amplifier of two stages is shown in the illustration below. Here a DC voltage is applied directly to the grid of the first stage. The output of the plate of this stage is connected directly to the grid of the second stage. The bias and B+ voltage on each succeeding stage of a direct-coupled amplifier must be greater than on the preceeding one. Direct-coupled amplifiers are used in vacuum-tube voltmeters and electronic voltage regulators where small changes of DC voltage must be amplified.

When one audio (or radio) frequency amplifier stage is coupled to another, the plate of the first stage cannot be directly coupled to the grid of the following stage because of the difference in DC potential between the two. A coupling circuit must therefore transfer the varying signal voltage (AC) from one stage to the other and, at the same time, keep the different DC voltages from affecting each other. This double function is effectively accomplished by resistance-capacitance (RC), impedance, and transformer coupling circuits.

A *resistance-capacitance* (RC) coupling circuit between two amplifier stages is shown schematically at the top of page 221. This is the method used most widely for coupling audio stages. The bias for each stage shown is obtained through cathode bias. The capacitor C_c, the *coupling capacitor*, provides an AC current path to the grid of the following stage. The load resistor R_L is a resistance of large value so that a large voltage drop can be coupled to the grid of the next stage. When a signal



A Direct-coupled Amplifier



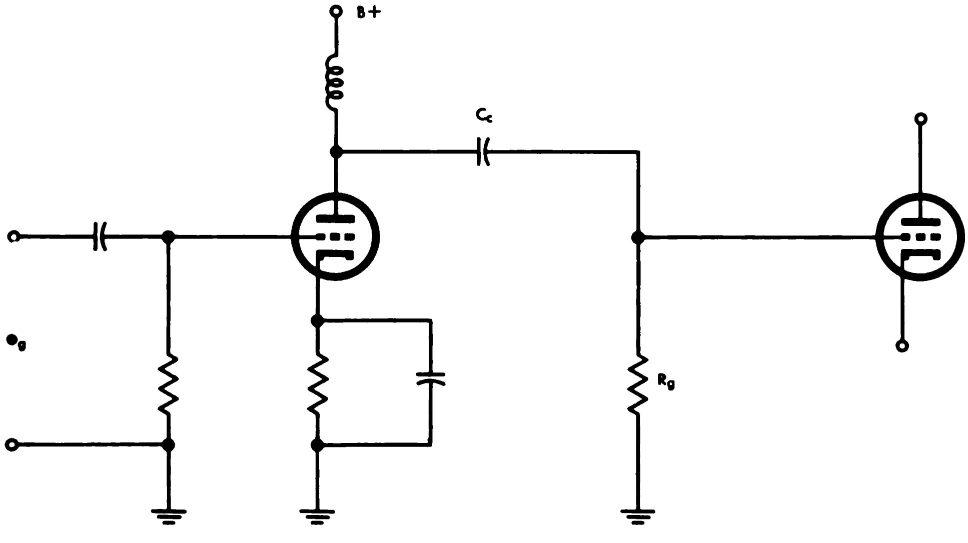
RC-Coupled Amplifier

is applied to the grid of the first stage, voltage variations occur across R_L. These voltage variations are coupled to the grid of the second tube through the RC circuit composed of C_c and R_g, the grid resistor. The coupling capacitor performs the double function of *providing a low-impedance path for the AC* and of *blocking the DC plate voltage from the grid of the following tube.*

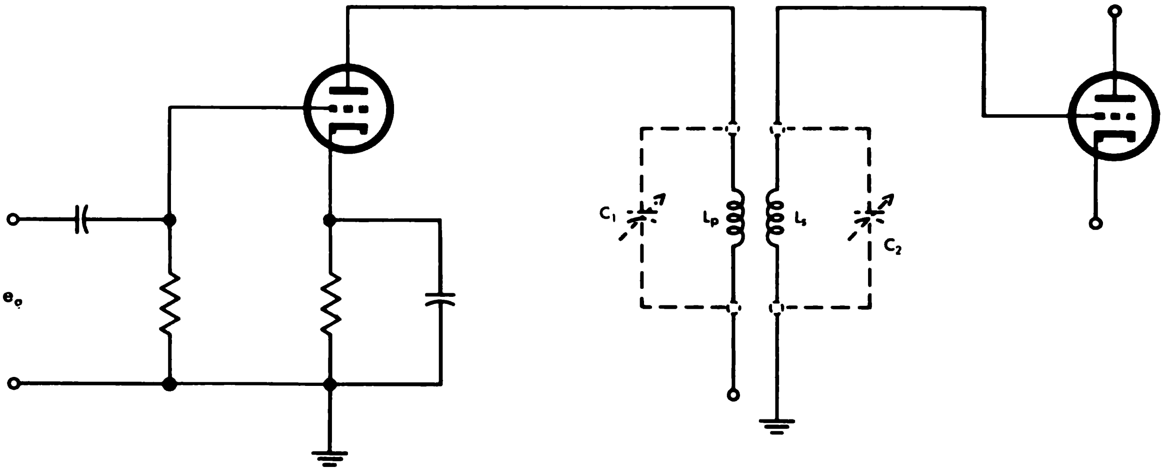
An *impedance-coupling* circuit is shown in the schematic below. The circuit differs from the RC coupling circuit in that an inductance L replaces the resistor R_L in the plate circuit of the RC type. This inductance makes possible a greater gain for a given higher band of frequencies. This greater gain is due to the fact that the impedance of the inductance is very low for DC and increases as the frequency of the signal increases. Before a signal is applied, the DC plate voltage is higher if the plate load is an inductance rather than a resistance. For a given band of high frequencies, the impedance of L becomes greater than that of R_L in the RC-coupled amplifier. Therefore, a greater variation occurs across the inductance at these higher frequencies, and a greater gain results.

Transformer coupling is another type of coupling you will often encounter. An example of transformer coupling is shown in the illustration on page 222. An untuned transformer is often used to couple audio frequencies. Because of the step-up property of the transformer, greater gain can be obtained with transformer coupling than with RC or impedance coupling. But this greater gain is obtained at the expense of narrowing the bandwidth that is amplified. Note that the transformer also performs the double function of coupling the AC variations and blocking the DC plate voltage from the grid of the next stage.

The tuned transformer is generally used to couple radio (including intermediate) frequency amplifier stages. Capacitors (shown dotted) are placed in parallel with the primary and secondary of the transformer, forming two parallel-tuned circuits. Greater gain with a narrower bandwidth results from tuned transformer coupling if the Q of each of the tuned circuits is high.



Impedance-Coupled Amplifier



Transformer Coupling

DIRECT CURRENT AMPLIFIERS

As stated earlier in this chapter, amplifiers are classified according to the band of frequencies they amplify, as follows: direct-current amplifiers, audio amplifiers, video amplifiers, and radio-frequency amplifiers.

Direct-current amplifiers are used to amplify very-low-frequency or DC voltages. A simple DC amplifier consists of a single tube with a grid resistor across the input terminals and with the load in the plate circuit. The load may be some sort of mechanical device, such as a relay or a meter, or the ouput voltage may be used to control the gain of an amplifier.

The DC voltage that is to be amplified must be applied directly to the grid of the amplifier tube. For this reason, only direct coupling can be used in the amplifier input circuit. Examples of DC amplifiers are given on page 220 under the discussion of direct coupling.

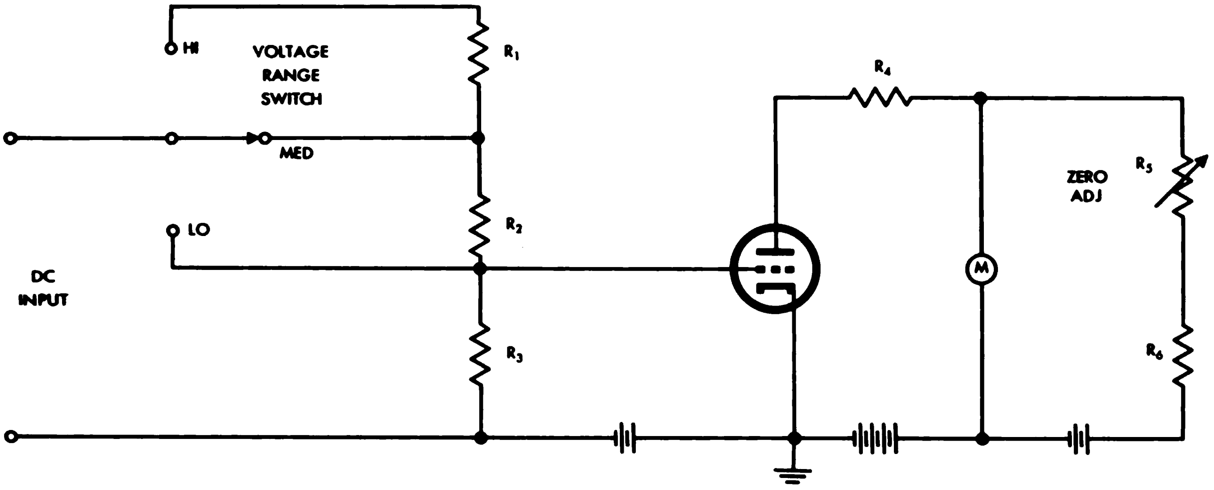
One of the most important uses of a DC amplifier is in a vacuum tube voltmeter. Study the circuit of the vacuum tube voltmeter shown on page 223. The voltage to be measured is applied to the voltage divider, made up of R_1, R_2 and R_3. The ratio of voltage division can be varied by a voltage range switch in such a way that several ranges of voltage can be measured. Resistor R_4 is used to prevent damage to the meter if too high a voltage is applied. In the plate circuit, the additional battery and the variable resistor are used to balance the normal plate current of the circuit. The variable resistor can be adjusted so that Meter M reads zero when no signal is applied. Whenever a DC voltage is applied to the input, the tube amplifies it and causes a current to flow through the meter. Since the meter reading is proportional to the voltage applied, you can read the amount of voltage on the calibrated scale of the meter.

DC amplifiers are also used in electronic voltage regulators where small changes of voltage must be amplified for control purposes. Another important application of DC amplifiers is found in the automatic frequency control (AFC) circuits of many radar sets. In these circuits a frequency change is converted to a voltage change and then amplified by a DC amplifier. AFC circuits are discussed later in this manual.

AUDIO AMPLIFIERS

Amplifiers designed for frequencies within the audio spectrum (approximately 20 to 20,000 cycles) are called audio frequency amplifiers.

To obtain sufficient gain, it is usually necessary to use one or more stages of audio amplification — that is, to use the output of one tube to control the grid circuit of a second, then to use the amplified output of the second tube to control the grid of the following tube, and so on. When two or more amplifiers are connected in this manner, they are said to be connected in *cascade*.

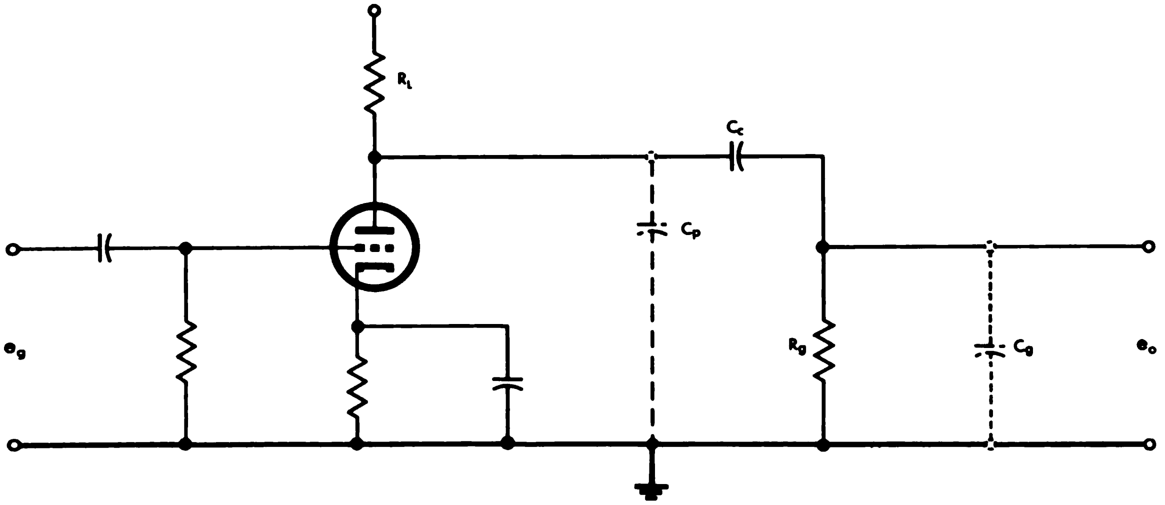


Vacuum Tube Voltmeter

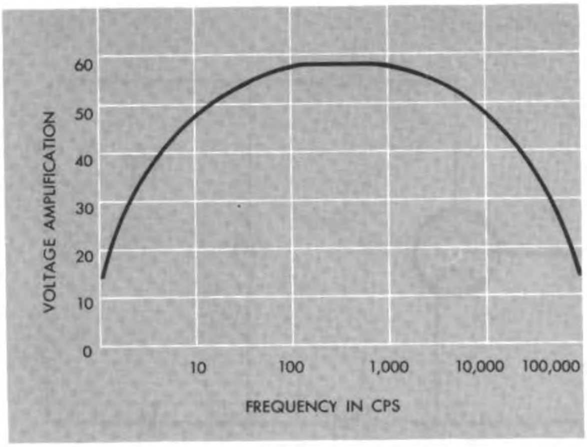
RC coupling is the most widely used type of coupling in audio circuits; however, impedance and transformer are also found. The circuit shown below is a typical RC-coupled, audio amplifier. The capacitors drawn dotted represent the interelectrode and stray wire capacitances present in the plate circuit of one tube and the grid circuit of the following stage.

The amplification of an RC-coupled, audio amplifier varies with the frequency of the input signal. The first frequency response curve shown on page 224 represents this variation in amplification over a wide range of frequencies. Note that the gain of the amplifier decreases at the low and the high frequencies and remains fairly constant in the middle range of frequencies. The low response at low frequencies is caused by the large reactance of the coupling capacitor. At the high frequencies the low response is due to the shunting effect of the interelectrode and stray wire capacitances of the plate and grid circuits. Both these effects are negligible in the middle range of frequencies.

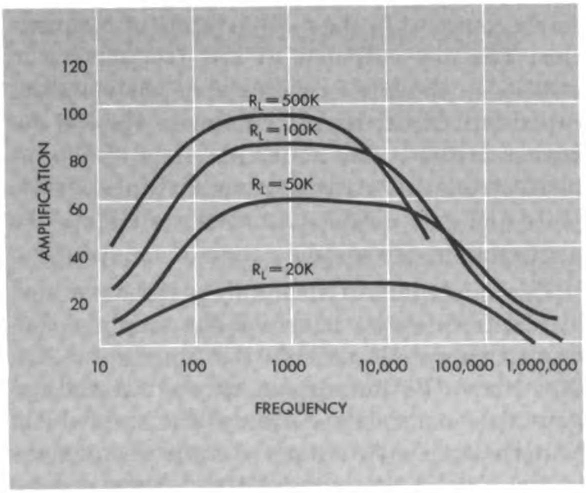
The bandwidth of an amplifier is determined by the frequencies at which the amplification drops to approximately 70% of maximum. On this curve the line drawn through a voltage amplification of 42 determines the bandwidth. Note that the entire range of audio frequencies is included in the bandwidth.



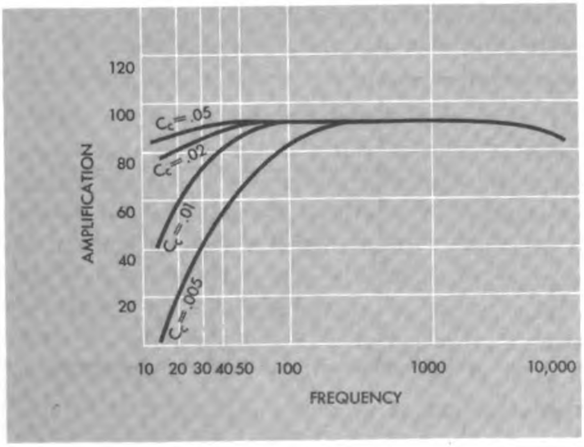
RC-Coupled, Audio Amplifier



Frequency Response of RC-Coupled Audio Amplifier



Effect of R_L on Frequency Response



Effect of Coupling Capacitor on Low Frequency

VIDEO AMPLIFIERS

Radar and television signals require amplifiers that can amplify a wide band of frequencies. Amplifiers that have a uniform response over a band of frequencies extending from the lowest audio frequencies to frequencies of several megacycles per second are known as video amplifiers.

RC coupling is used in most video circuits; the values of circuit components, however, are different from those used in audio circuits. The frequency response curve of the RC-coupled audio amplifier shown to the left drops rapidly at the low and high frequencies but is fairly uniform over the middle range of audio frequencies. The low-frequency response is limited by the value of the coupling capacitor, and the high-frequency response is limited by the interelectrode and stray wire capacitances of the plate and grid circuit. To widen the bandpass of this amplifier so that its response is uniform over the wide range of frequencies required by a video signal, steps must be taken to offset these limiting factors.

Triodes, which have high interelectrode capacitances, are not suitable for use as video amplifiers. Therefore, pentodes, which have very low interelectrode capacitance, are used. In addition, pentodes also have the advantage of higher amplification than the triode. Care must be taken in wiring the circuit to keep leads short and properly spaced to reduce to a minimum the distributed capacitances in the wiring.

The value of the load resistor in RC-coupled amplifiers determines to a great extent the bandwidth and gain of the amplifier. The effect of changing the value of load resistor R_L in the circuit is shown in the chart to the left. Note that the resistor of high value produces a high output voltage in the middle-frequency range and a steep drop in output voltage at the high frequencies. As the value of R_L is decreased, the response is more uniform over a wider range, although the gain in the middle-frequency range is lower.

The amplification at the high frequency end is limited by capacitances that shunt the load resistor. These capacitances, which are dotted and labeled C_p and C_g in the circuit illustration, consist of the wiring and interelectrode capacitances of the tubes. C_p acts as a low impedance in parallel with R_L and reduces the gain of the stage at the higher frequencies. Video amplifiers, which must have a fairly uniform response over a wide band of frequencies, employ pentodes to reduce the interelectrode capacitances and load resistors of low value to broaden the frequency response. In addition, special devices often are added to improve the response at the low frequencies (low-frequency compensation) and at the high frequencies (high-frequency compensation).

Low Frequency Compensation

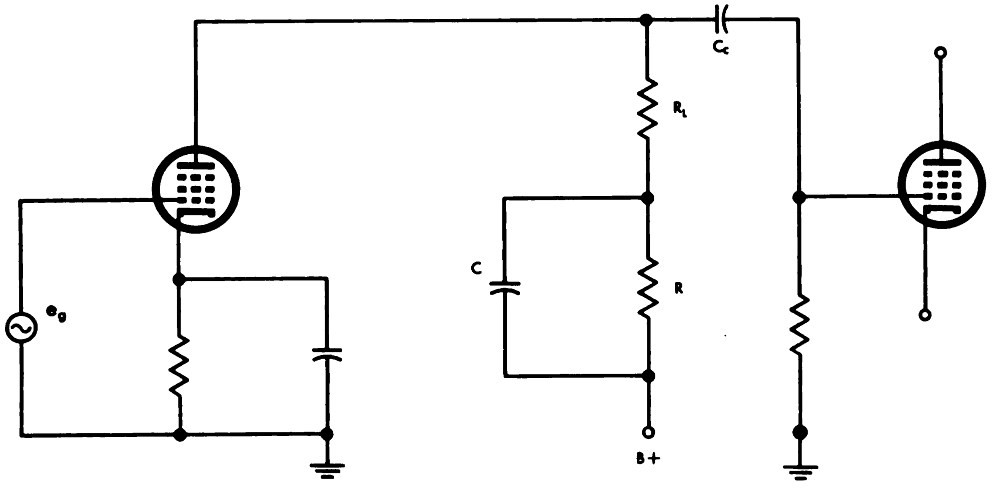
Video amplifiers must amplify frequencies at the low end of the bands as well as at the high end. According to the chart shown at the bottom of page 224, amplification of low frequencies shows improvement when the coupling capacitor C_C is large. A large capacitor has less reactance to low frequencies than a small one, and low reactance means a smaller voltage drop. Since the coupling capacitor and the grid resistor form a voltage divider in the input circuit, more voltage appears across the grid resistor when C_C has little reactance. Therefore greater amplification takes place in the tube. There is a limit, however, to the increase in amplification possible by increasing the size of C_C, particularly in radar. If C_C is too large, it discharges too slowly for the faithful reproduction of the steep trailing edges found in radar waveshapes.

A more satisfactory means for improving low frequency amplification is to use a low frequency compensating circuit similar to the one shown below. A low frequency compensating circuit improves response by using a parallel RC circuit in series with the load resistor. Since C is comparatively large, it offers practically no reactance to middle and high frequencies and, therefore, does not affect these frequencies. At low frequencies, however, its reactance is high. This reactance in parallel with R produces an impedance which, when added to R_L, increases the total load impedance. Since the compensating circuit produces a larger load impedance at low frequencies, the gain of the stage is higher at these frequencies.

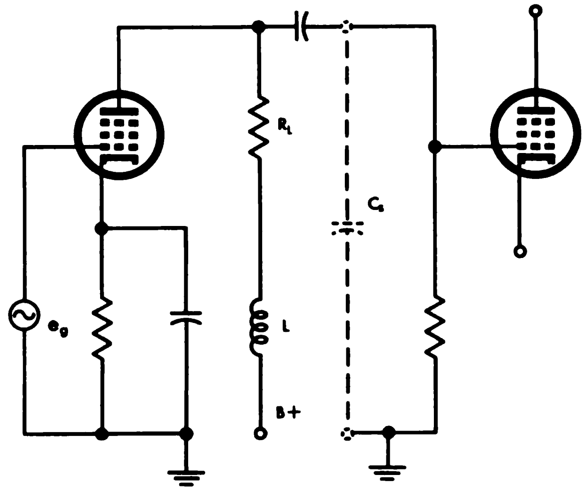
The low frequency response of a video amplifier may therefore be improved by using a coupling capacitor of large value in combination with a low frequency compensating circuit in the plate load. If the values of the capacitors are correctly chosen, the increased voltage drop across C_C at the low frequencies is compensated for by the increased amplification caused by the added load impedance.

High Frequency Compensation

Although the bandwidth of an amplifier is increased by using load resistors of low value, the amplification at the high frequency end is limited by the capacitances (C_s) which shunt the

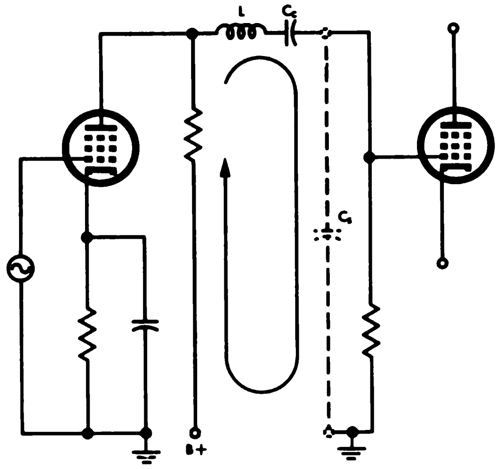


Low Frequency Compensation Circuit



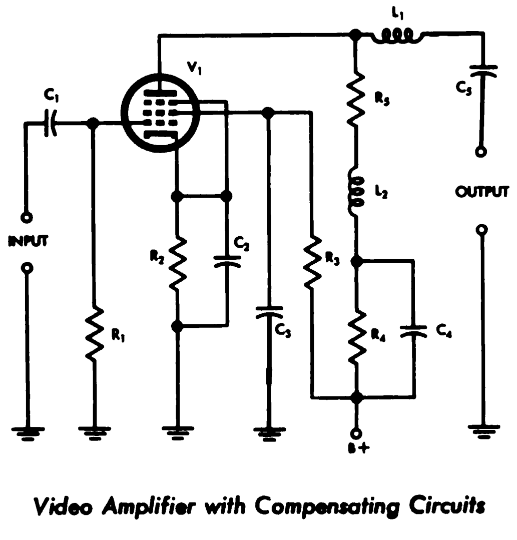
Shunt Peaking Coil for High Frequency Compensation

load resistor. High frequency compensating circuits are used in many video amplifiers to offset the effects of the shunt capacitances. One such circuit uses a ***shunt peaking*** coil. As shown in the circuit to the left, a low reactance coil (L) is placed in series with the load resistor (R_L) and in shunt with C_s. At low and middle frequencies, the reactance of the coil is low. However, at high frequencies, the reactance is high and the gain of the amplifier increases. This increase in gain offsets the loss caused by the shunting effect of C_s. The value of the coil is usually chosen so that it forms a resonant circuit with the shunt capacitances at a frequency where the gain is low and increases the gain at this frequency. In this way, the shunt peaking coil extends the band of frequencies which the video amplifier can amplify uniformly.



Series Peaking Coil for High Frequency Compensation

The series peaking coil is another type of high frequency compensating circuit. As shown in the illustration to the left, a coil L is connected in series with the coupling capacitor C_c. At high frequencies, the inductance L, capacitor C_C, and capacitor C_s form a series resonant circuit. The increased current in this resonant circuit causes an increased voltage drop across C_s, and a greater voltage is applied to the grid of the following stage.

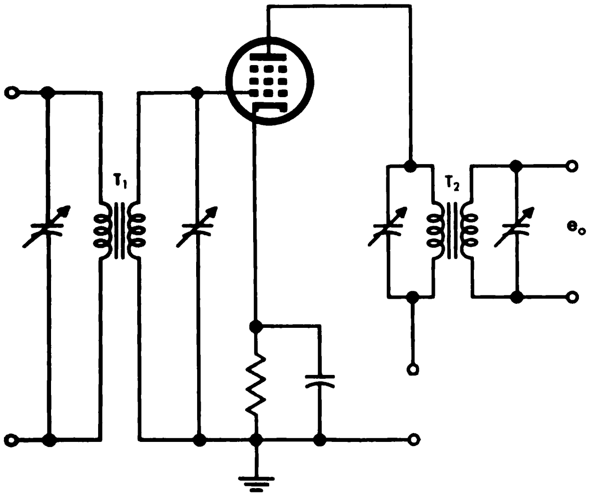


Video Amplifier with Compensating Circuits

The two methods of high frequency compensation are often combined in video amplifiers to increase the bandwidth at the high frequency end. The schematic drawing to the left is an RC-coupled amplifier in which are incorporated all the devices for extending the bandwidth so that it can function as a video amplifier. Study this circuit and note the components which extend the bandwidth at the low and high frequency ends.

RF AMPLIFIERS

Radio frequency amplifiers amplify voltage or power at radio frequencies. They are used in transmitters and receivers. Intermediate-frequency amplifiers are a special type of RF amplifier used in superheterodyne receivers. IF and RF amplifiers are essentially alike. They differ only in their use and in the radio frequency they amplify. For example, a typical RF amplifier for a radio receiver might amplify a frequency of 610 kc while the IF amplifier would be tuned to 455 kc. A radar receiver, on the



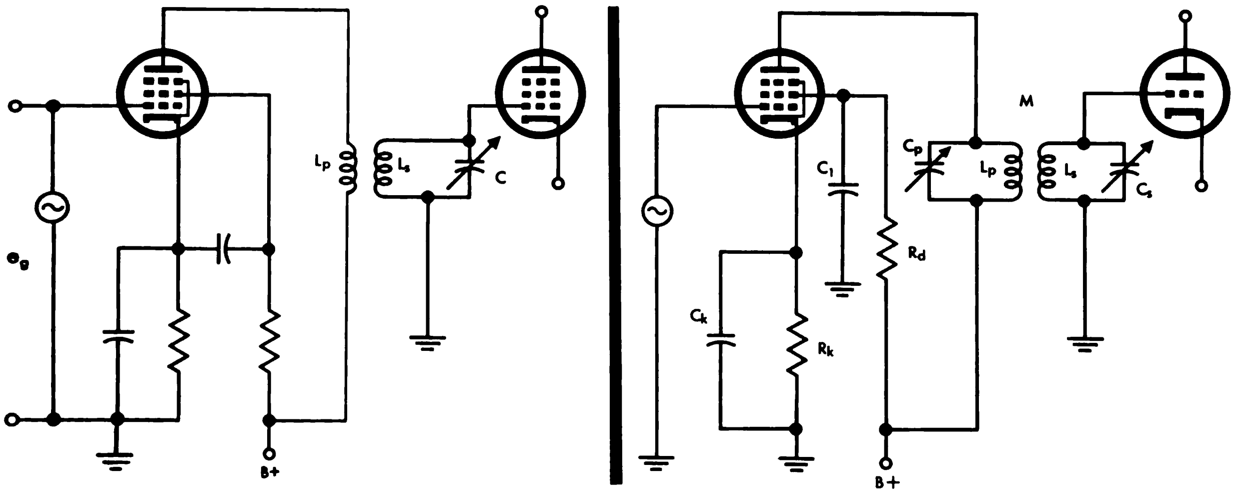
RF Amplifier

other hand, might have RF amplifiers at 175 megacycles and IF amplifiers at 30 megacycles. Radar receivers above 1,000 megacycles do not use RF amplifiers; however, they use many stages of IF amplification.

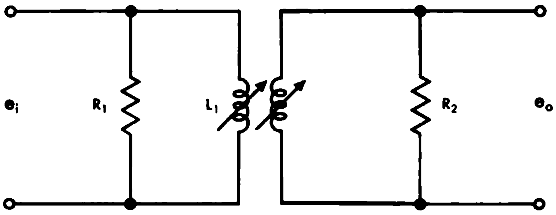
RF amplifiers, like audio amplifiers, may operate class A, B, or C. They differ from audio amplifiers in that resonant circuits are usually used for input and output coupling circuits. For low-power operation, such as that required in receivers, pentode RF amplifiers provide the highest gain, and, because of their low grid-to-plate capacity, have the least tendency to break into self-oscillation. The RF amplifier shown at left is actually an IF amplifier, since it is designed to amplify only one frequency. Single frequency amplification occurs because of the parallel resonant circuits at the input and output ends. These parallel circuits discriminate against all other frequencies by developing maximum voltage at the resonant frequency, and minimum voltage at other frequencies. This ability of an amplifier to select one frequency and reject others is called *selectivity.* Selectivity in an amplifier is greatest when high Q, tuned circuits are used.

IF amplifiers used in radar receivers must be of special design. A pulse modulated RF signal contains the radio frequency and sidebands for each of the frequencies which make up the pulse. The IF amplifier must have a very wide bandwidth to faithfully amplify these input pulses. The devices used in radar IF stages to increase the bandwidth to pass a band from three to six megacycles reduces the gain of each stage. Radar receivers, therefore, usually have six or more IF stages.

IF amplifiers may be *single-tuned* or *double-tuned.* As you can see in the illustration below, a single-tuned stage has one tuned circuit between input and output stages, while the double-tuned stage has two. In each case, the transformer is tuned to resonance with the stray and interelectrode shunt capacitances. Double-tuned coupling produces greater gain for a given bandwidth, but single-tuned coupling is simpler to align and manufacture.



Single-tuned and Double-tuned IF Amplifiers



Lowering the Q of Tuned Circuits

IF bandwidths may be increased in any of the following ways:

1. By *lowering the Q of the tuned circuits.* This may be done by resistive loading of the tuned transformer. The smaller the value of the resistors R_1 and R_2 in the illustration, the wider is the bandwidth and the lower is the gain.
2. By *using stagger tuning*. This means tuning some IF stages slightly below, some slightly above, and some to the intermediate frequency. The illustration below shows stagger tuning between IF stages. This method of tuning stages increases the bandwidth and lowers the over-all gain.
3. By *using close coupling* in double-tuned stages.
4. By *using an unbypassed cathode resistor*.

GAIN AND VOLUME CONTROLS

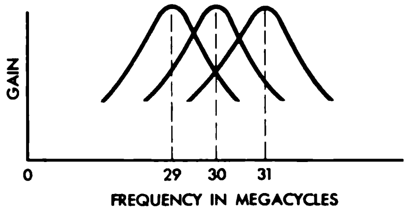
The volume control of the radio receiver and the gain control of the radar receiver are alike in the function they perform. Both controls determine the strength of the output signal which is reproduced either by a speaker or cathode-ray tube.

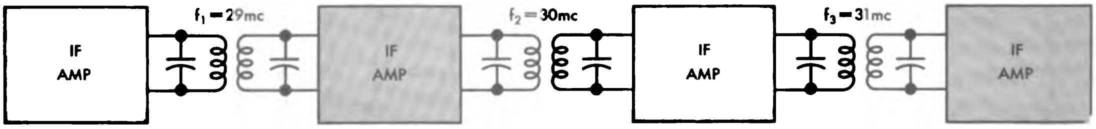
Volume Control

The volume control usually performs its function by the selection of a portion of an available signal which is to be amplified by the amplifier stages. Two variations of manual volume controls are shown in the illustration at the top of the next page. Note that only a portion of the input signal is applied to the control grid of the audio amplifier in each case. At A, the input signal is capacitively coupled to a potentiometer. The movable tap on the potentiometer is connected directly to the grid. The potentiometer acts as the grid resistor, while the tap selects the portion of the signal to be amplified. At B, transformer coupling is being used. Here the potentiometer is across the secondary of the transformer and the movable tap is connected directly to the grid.

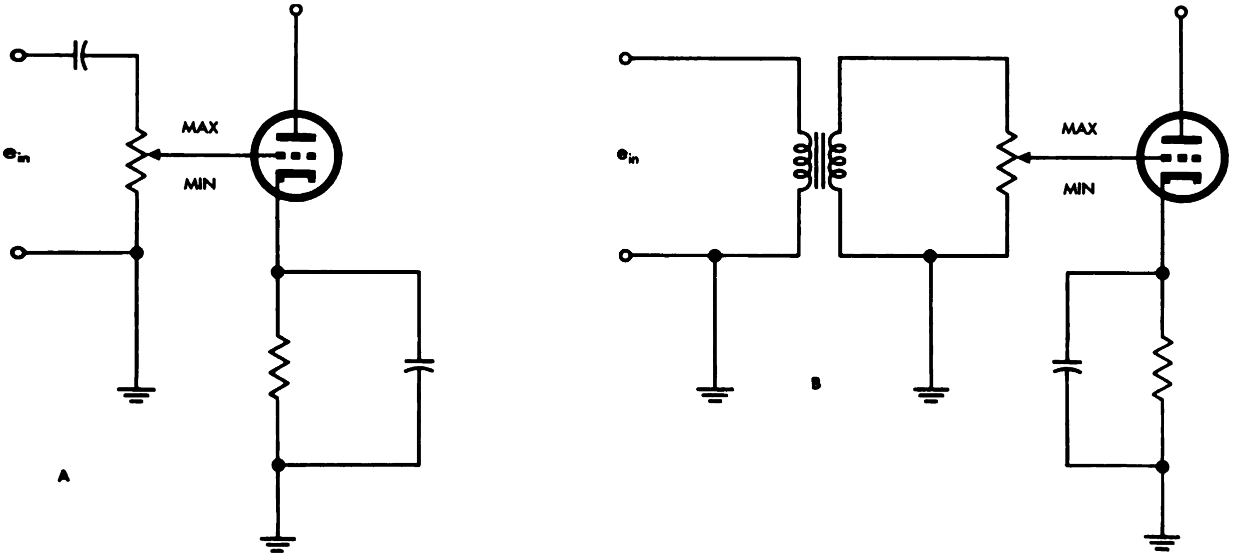
Gain Control

The gain control performs its function by regulating the amount of amplification of a stage or stages. As you learned in your study of remote-cutoff tubes, the gain of a remote-cutoff tube may be varied by varying the bias. Therefore, a control which varies the bias of a remote-cutoff tube can serve as a gain control. The illustration at the bottom of the next page is an example of this type of manual gain control circuit. Here the potentiometer in the cathode circuit is in series with a fixed resistor which provides cathode bias for the tube. Adjusting the potentiometer varies the bias and, therefore, the amplification of the stage. In actual use, the potentiometer is in the cathode circuit of several RF stages and controls the gain of these stages.





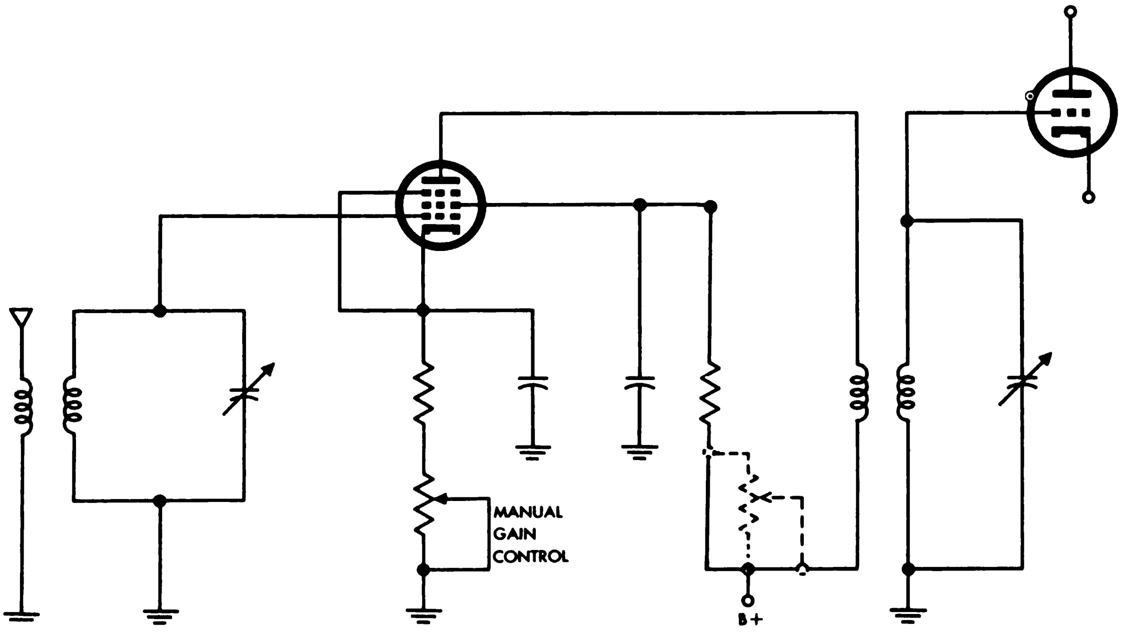
Stagger Tuning



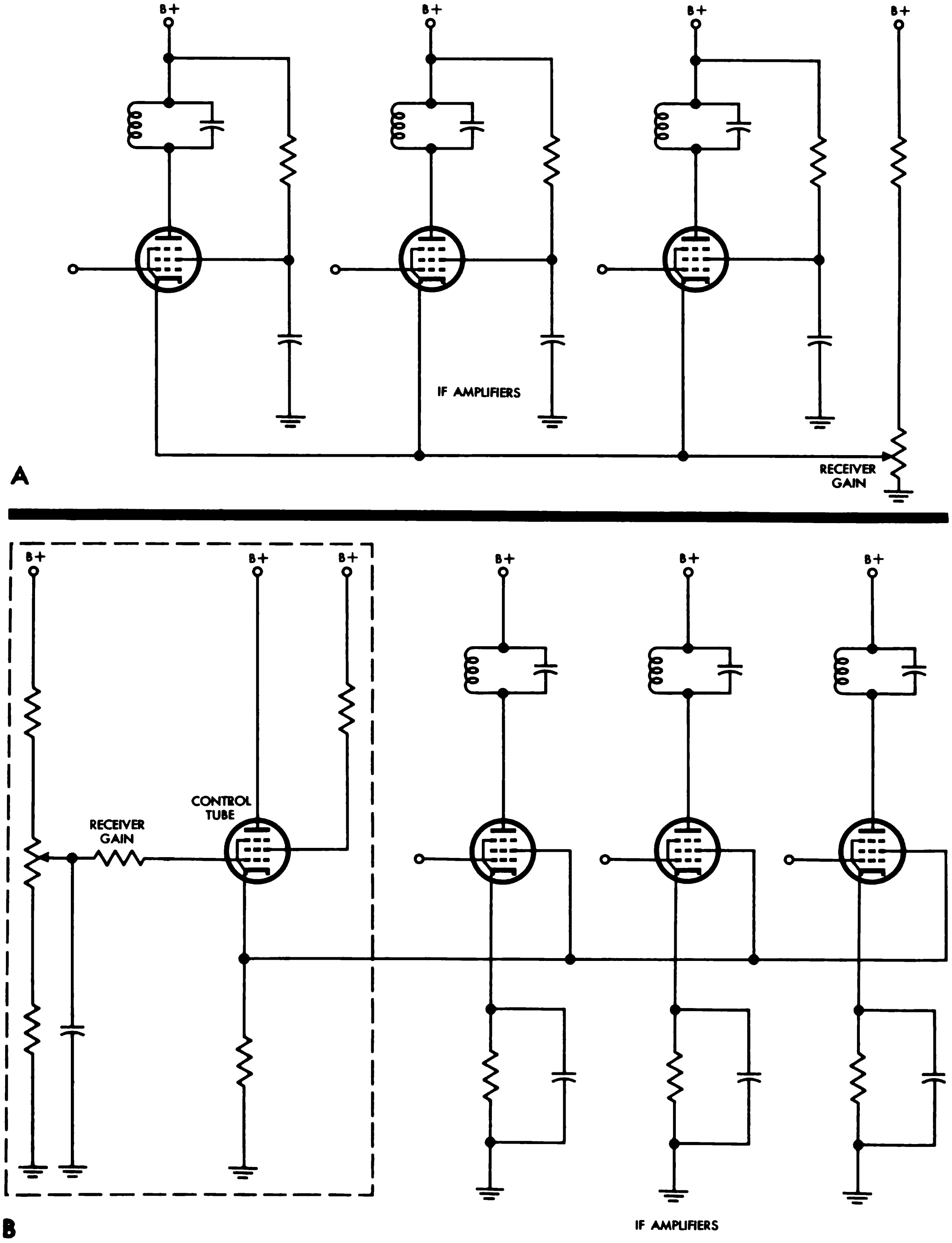
Manual Volume Controls

Another method of changing the amplification of a pentode is to vary the voltage between the screen grid and cathode of the stage. The potentiometer shown dotted in the illustration is an example of this type of gain control. As the voltage applied to the screen grid of the tube is increased, the gain is increased.

Both types of gain controls are used in radar receivers. Two typical gain control circuits are shown in the illustrations on page 230. At A, the receiver gain control varies the bias applied to the cathodes of three IF stages. Since these tubes are remote-cutoff tubes, the gain of these stages varies with the change in bias. At B,



Manual Gain Control Circuit



Receiver Gain Control Circuits

the receiver gain control varies the bias on the grid of the control tube. This, in turn, varies the current which flows through the control tube and the voltage across its cathode. Since the cathode is connected to the screen grids of the three IF stages, the voltage change varies the gain of these stages.

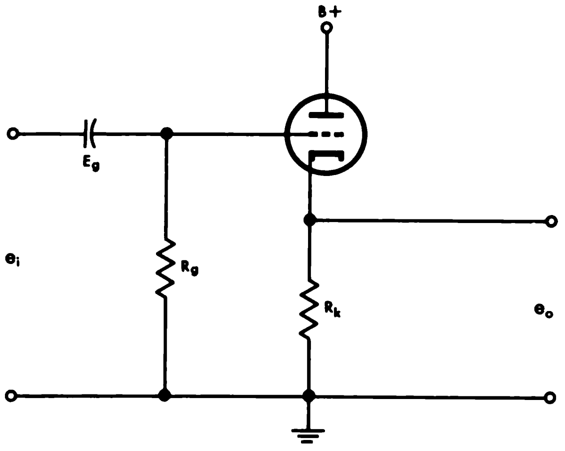
CATHODE FOLLOWER

We learned in our study of transformers that maximum transfer of power takes place when the impedance of the source is matched to the impedance of the load device, and that a transformer is frequently used as a means of impedance matching. The transformer is limited, however, in its use because of its comparatively poor frequency-response characteristics. Where fidelity is of prime importance, as in radar, a vacuum-tube circuit is commonly used to match unequal impedances. Since the problem of impedance matching is usually one of connecting the output of a high-impedance source to a low-impedance device (such as a coaxial cable), the impedance-matching circuit must have a high-impedance input and a low-impedance output. A circuit which meets these requirements is the cathode follower.

A cathode follower is shown schematically in the illustration to the right. The input is applied to the grid of the tube, and the output is taken across the cathode resistor. Since the grid resistor (R_g) is of large value and the cathode resistor of small value, the circuit fulfills the requirement of high input impedance and low output impedance.

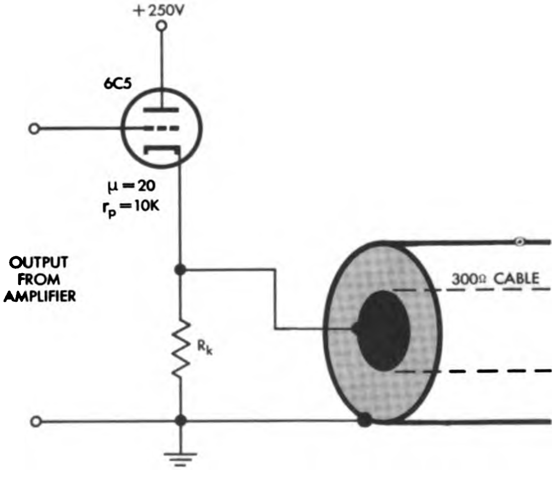
The cathode follower is an amplifier in which inverse feedback occurs because of an unbypassed cathode resistor, across which the output is taken. The circuit is essentially an impedance-matching device with a voltage gain less than unity. It does, however, yield a power gain and is therefore classified as a power amplifier. The frequency response of a cathode follower is excellent; it is often used as the final video stage of a radar receiver.

In radar, the output from one unit of equipment is often fed through a long cable to another unit. Such a cable, because of its distributed capacitance, inductance and resistance, presents an impedance which is much lower than the output impedance of the ordinary amplifier. A cathode follower, on the other hand, may be designed to have a very low output impedance. The use of a cathode follower instead of a conventional amplifier as shown below, therefore, permits the transfer of much more power from one unit to another.



A Cathode Follower

As the input voltage to a cathode follower goes more positive, the plate current increases. The cathode-to-ground voltage therefore goes more positive. As the input voltage goes more negative, the plate current decreases and the cathode-to-ground voltage decreases. Note that the cathode voltage *follows* the input signal; hence the name, cathode follower.



Cathode Follower with Coaxial Cable as Load Impedance

Voltage Gain

The voltage gain (A) of a cathode follower is:

$$A = \frac{\mu R_k}{r_p + R_k (\mu + 1)}$$

where μ is the amplification factor of the tube, r_p the AC plate resistance, and R_k the cathode resistor. Since the denominator of the equation is greater than the numerator, the gain of a cathode follower is less than one. Remember, though, that the gain is less than one only for AC signal voltages; the cathode-to-ground DC voltage is greater than the input voltage.

Output and Input Impedances

The input impedance of a cathode follower is high. Since a cathode follower is operated with the grid negative with respect to the cathode, a high amplitude voltage can be applied between the grid and ground without causing grid current flow. This is due to the degenerative action of the unbypassed cathode resistor and the high input impedance of the cathode follower to a positive input signal. Since a cathode follower is a degenerative circuit, very little amplitude distortion is introduced in the output.

The output impedance (Z_o) of a cathode follower is low and is expressed as:

$$Z_o = \frac{r_p R_k}{r_p + R_k (\mu + 1)}$$

For a high-mu tube,

$$Z_o = \frac{R_k}{1 + R_k gm}$$

The equation shows that the output impedance of a cathode follower is always less than the impedance of the cathode resistor.

Cathode Resistor

You can see from the equation for the output impedance of a cathode follower that by choosing the proper tube and the proper size cathode resistor, you can match the low output impedance of the stage to a low-impedance load. The value of the cathode resistor is an important design consideration in that it largely determines the output impedance.

To find the proper value of cathode resistor to match a given output impedance, use the following equation:

$$R_k = \frac{Z_o r_p}{r_p - Z_o (\mu + 1)}$$

where Z_o equals the output impedance, r_p the plate resistance, and μ is the amplification factor of the tube.

Using the values in the lower figure on page 231, find the proper size of cathode resistor to match the impedance of the 300-ohm coaxial cable, and find the gain of the stage.

To find R_k, substitute in the formula

$$\begin{aligned} R_k &= \frac{Z_o\, r_p}{r_p - Z_o (\mu + 1)} \\ &= \frac{300 \times 10^4}{10^4 - 300\ (20 + 1)} \\ &= \frac{3{,}000{,}000}{10{,}000 - 6300} \\ &= \frac{3{,}000{,}000}{3700} \\ R_k &= 811 \text{ ohms (approximately)} \end{aligned}$$

To find the gain, substitute in the formula

$$\begin{aligned} A &= \frac{\mu R_k}{r_p + R_k (\mu + 1)} \\ &= \frac{20 \times 811}{10^4 + 811\ (20 + 1)} \\ &= \frac{16{,}220}{10^4 + 17{,}031} \\ &= \frac{16{,}220}{27{,}031} \\ A &= 0.6 \text{ (approximately)} \end{aligned}$$

The characteristics of the cathode follower may be summarized as follows:

1. It has a high input impedance and low output impedance and is, therefore, used to match a high-impedance source to a low-impedance load.
2. The gain is always less than 1.
3. The high-frequency response is excellent because the output impedance is low.
4. There is very little distortion of the signal due to the degenerative feedback in the circuit.
5. The output and input signals are always in phase.

Chapter 27

Principles of Oscillation

You have seen that electron tubes can amplify a signal. This ability of the electron tube to amplify also makes possible the generation of alternating currents. When used for this purpose, the electron tube and its associated circuit is called an *oscillator*.

Oscillators are of many types and perform many functions. The frequencies of the alternating currents generated by an electron tube oscillator may be only a few cycles per second or thousands of megacycles per second, depending upon its purpose. Oscillators produce the RF energy required in radio, radar, and television transmission. An oscillator is also an essential part of the radio, radar, or television receiver used for the reception of these signals. In addition, oscillators have numerous other applications. Their use in test equipment of various types has improved the servicing and repair of the most complicated electronic equipment. Induction heating, sorting, sampling, counting, and switching are a few examples of the countless functions performed by oscillators in the industrial field.

Let us see how the oscillator functions to generate alternating currents.

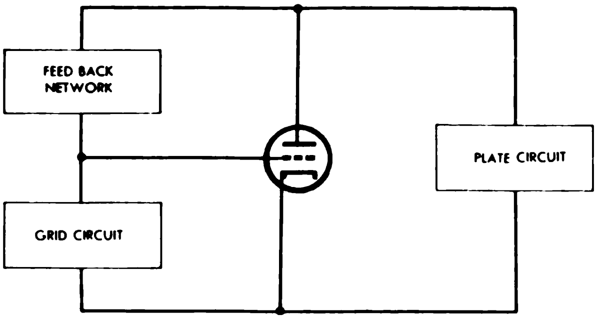
CHANGING AN AMPLIFIER TO AN OSCILLATOR

The amplifier, as you have already learned, takes an AC signal applied to its control grid and amplifies it in the plate circuit. This is accomplished by means of an electron tube, which converts DC power used in the plate circuit into an AC signal. The AC signal on the plate is an amplification of the input signal applied to the grid. The development of tubes with sufficient gain makes it possible to feed some of the signal from the plate circuit back to the input (grid) circuit. Feedback may either oppose or aid the input signal. If the feedback signal is of such phase as to oppose the input signal as in the cathode follower, the amplifier is called a *degenerative* or *negative* feedback amplifier. If the feedback signal is of such phase as to aid the input signal, the amplifier is called a *regenerative* or *positive* feedback amplifier. However, if the amplitude of a positive feedback signal is made so great that the input signal can be completely removed, and an AC output generated, the positive feedback amplifier becomes an oscillator.

In the illustration below, note that the oscillator consists of an electron tube which amplifies the signal on the grid and a feedback network through which part of the amplified signal is fed back from the plate to the grid. Each of these components introduces a phase shift of approximately 180°, causing a total phase shift of 360° from the grid to the plate and back to the grid. The feedback signal is therefore in phase with any signal which appears on the grid. Any slight change on the grid or plate is amplified and fed back to the grid in proper phase to reinforce the signal. The frequency at which oscillations are generated is determined by the circuit components. Once oscillations have started, the feedback voltage makes it possible for them to continue. You will see how this occurs as you read the sections which follow.

PRODUCING OSCILLATIONS IN A TANK CIRCUIT

A parallel resonant circuit is often called a *tank* circuit, because it has the ability to take energy fed to it in the form of pulses, store this energy alternately in the inductance and capacitance, and produce an output which is a continuous AC wave. You can understand how this is accomplished by a study of the illustration on page 235. When the switch is placed in position A, a pulse of DC voltage is applied across the capacitor. When the switch is moved to position B, the DC voltage is removed and the tank circuit is closed. The output across the tank circuit is an AC waveshape obtained as follows:



Block Diagram of an Oscillator

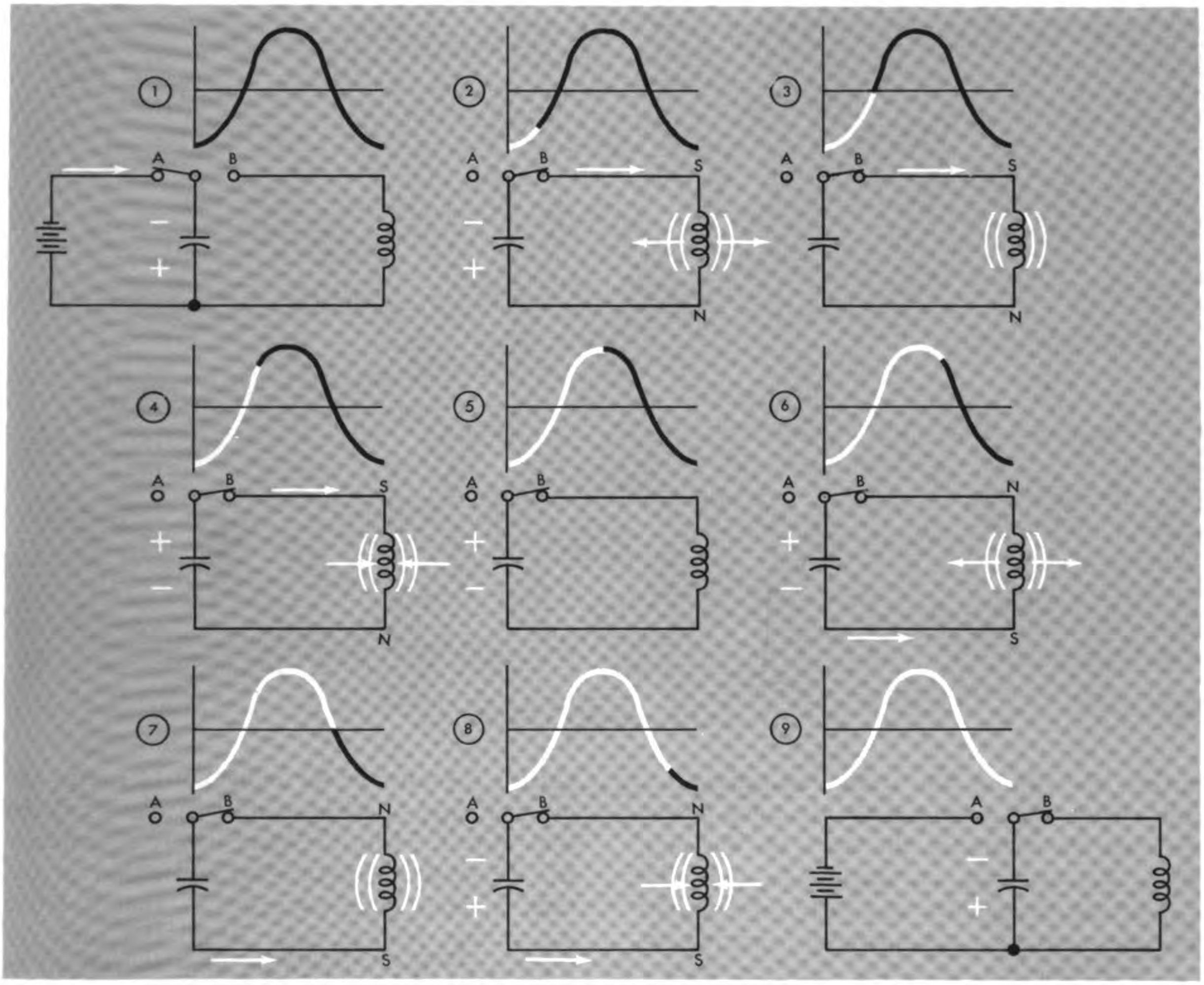
With the switch in position A, the capacitor charges quickly to the voltage of the battery with the polarity shown. As soon as the switch is placed in position B, the charged capacitor (which acts as a source) instantaneously begins to discharge through the inductance (which acts as a load). The current which flows through the coil as the capacitor discharges builds up a magnetic field around the coil (2). In this way, the energy stored in the dielectric is converted into energy stored in the magnetic field. The arrows shown are used to indicate the direction of the flow of electrons and the direction of the motion of the magnetic field. When the number of electrons on both plates of the capacitor is equal, no energy is stored in its dielectric and the current tends to stop flowing (3). The voltage across the capacitor is zero at this time.

NOTE

The white waveshape shown above each diagram represents the voltage between the *top* and *bottom* plates of the capacitor at that instant. It also represents the voltage across the coil and across the tank.

Since the flow of electrons tends to cease as the capacitor discharges, the inductance offers opposition to this change in the rate of flow of electrons. The collapse of the magnetic field (4) maintains the current flow in the same direction and charges the capacitor (which is now the load) with the polarity shown. This continues until the magnetic field completely collapses and the capacitor is charged (5). The output voltage of the tank circuit has a polarity which is now positive.

The circuit, however, is not balanced, for there is an excess of electrons on the bottom plate of the capacitor. The capacitor therefore begins to discharge again through the coil (6); however, the direction of the current flow is now reversed. The magnetic field builds up as the capacitor discharges, but the polarity of the output voltage remains positive while



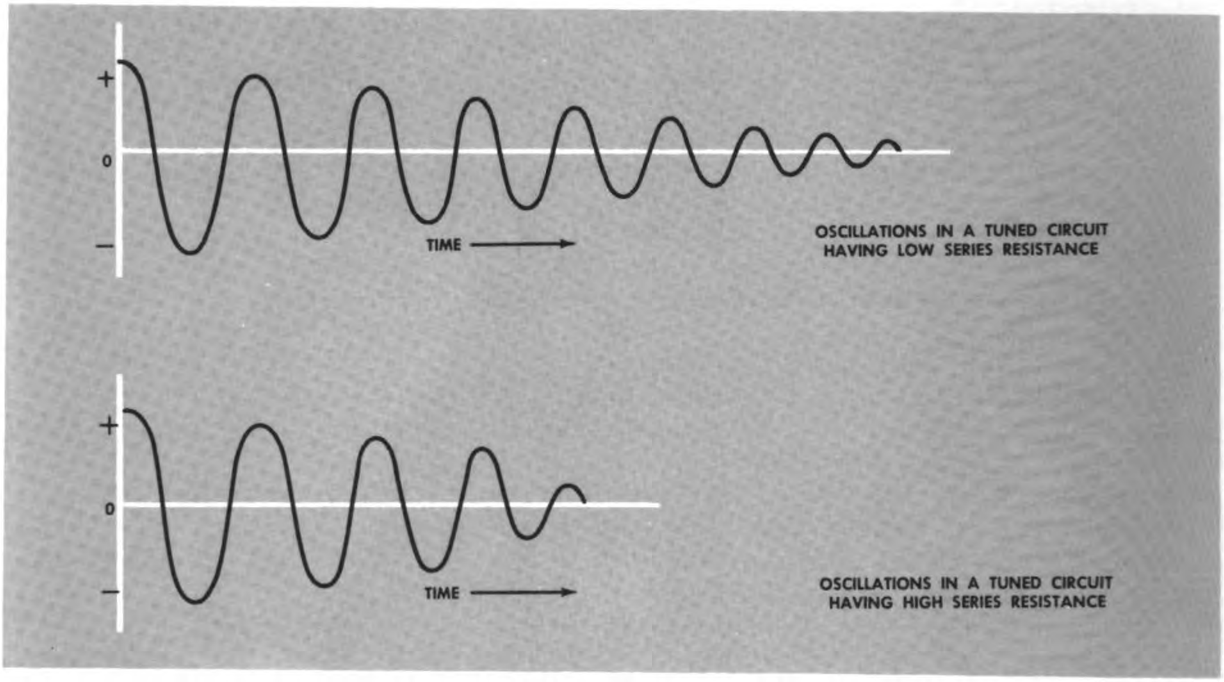
Producing Oscillations in a Tank Circuit

decreasing. This continues until the capacitor discharges and current flow tends to cease (7). At that instant, the magnetic field collapses, causing current to continue to flow in the same direction and charging the capacitor as shown (8). When the capacitor is charged, the conditions in the tank circuit have returned to those that existed at the instant the switch was placed in position B. The entire process (the discharge of the capacitor and the buildup of the magnetic field followed by the charge of the condenser and the collapse of the field) is repeated at a rate determined by the values of L and C. The output waveshape (9) taken across the tank circuit is a sine wave which occurs as the inductance and the capacitance alternately store and release the energy in their respective fields.

EFFECT OF RESISTANCE ON DAMPING

If there were no resistance in a tank circuit, oscillations would continue indefinitely without any decrease in amplitude. However, every resonant circuit contains some resistance; therefore, power is dissipated, and the energy alternately stored in the electrostatic and electromagnetic fields becomes progressively less. You can see this by examining the outputs of the resonant circuits shown in the illustration on the next page.

Note that resistance causes the amplitude of the oscillations to become progressively less and finally to cease completely. This effect on the output of a tank circuit, whereby the amplitude of each successive cycle becomes less, is known as *damping*. As you can see in the illus-



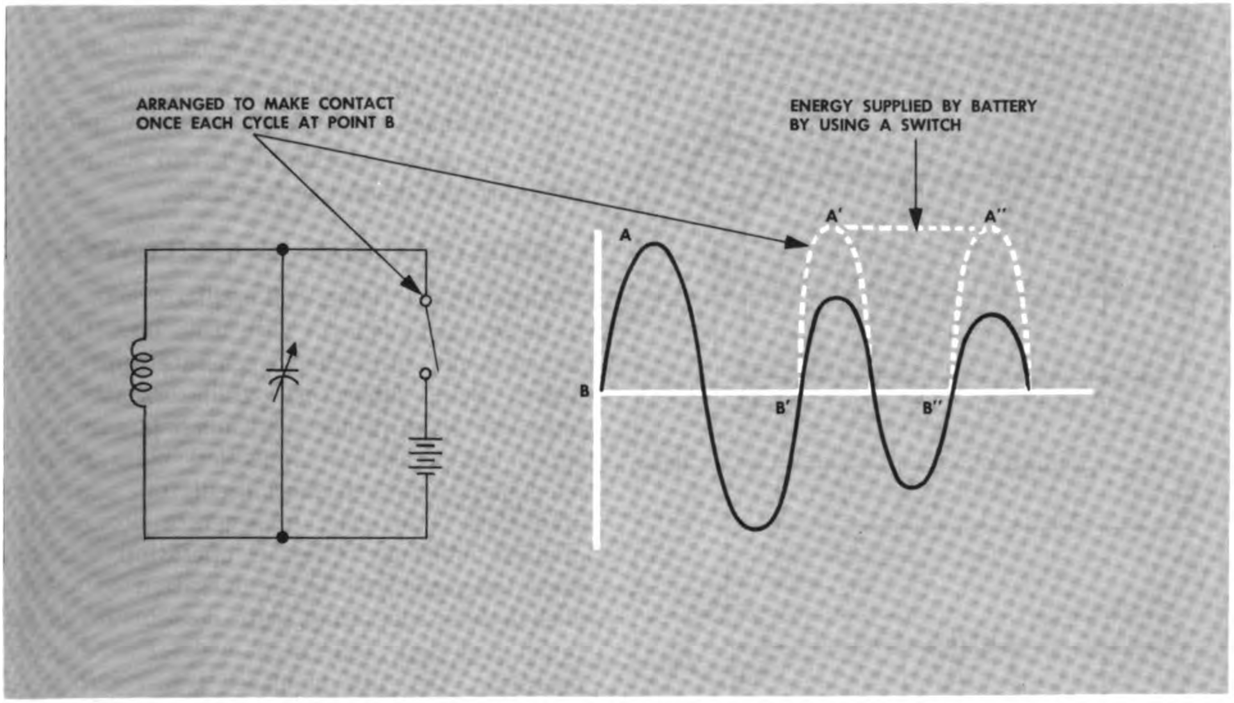
Effect of Resistance on Oscillations

tration, the greater the series resistance in a tank circuit, the greater is the damping effect; the lower the resistance, the less is the damping effect. You learned in chapter 18 that the Q of a tank circuit increases with a decrease in the series resistance of the circuit. A high Q circuit will therefore cause less damping than a low Q circuit.

The oscillations produced by charging the capacitor once and then allowing the oscillations to damp out are not usually desirable. In most cases, oscillations of constant amplitude are desired in the output. Since the damping effect is due to the loss of energy in the resistance contained in the circuit, either the complete removal of this resistance or the periodic replacement of the dissipated energy could sustain oscillations without damping. The complete removal of resistance is impossible in the construction of a practical circuit. Therefore, it is necessary to apply pulses of energy periodically to replace the dissipated energy, if undamped oscillations are to be sustained.

To see how this should work, consider the hypothetical situation of a tank circuit which is connected to, and disconnected from, a battery at an extremely rapid rate. In this setup, as shown in the illustration on page 237, the battery is connected for an instant to supply the pulse of energy needed to start oscillation in the tank circuit. Then it is immediately disconnected. When the first cycle of oscillation in the tank is about to be completed — that is, when the voltage reaches point B′ on the curve — the battery is connected for another instant to bring the voltage up to point A′, its original level. When the second cycle is about to be completed — that is, when the voltage reaches point B″ on the curve — the battery is connected for another instant to bring the voltage up to point A″, its original level. If the switching is handled quickly enough and if it is timed to occur at just the right point of each cycle, the output of the tank circuit will be oscillations of a constant amplitude.

The tank circuit, switch, and battery may be compared to a flywheel and a one-cylinder engine firing once each revolution. Just as the engine releases a pulse of mechanical energy once each revolution, the battery through the switch releases a pulse of electrical energy once each cycle. The flywheel stores and releases the mechanical energy, making the angular motion fairly uniform. The tank circuit stores and releases the electrical energy, making the voltage output a fairly good sine wave.

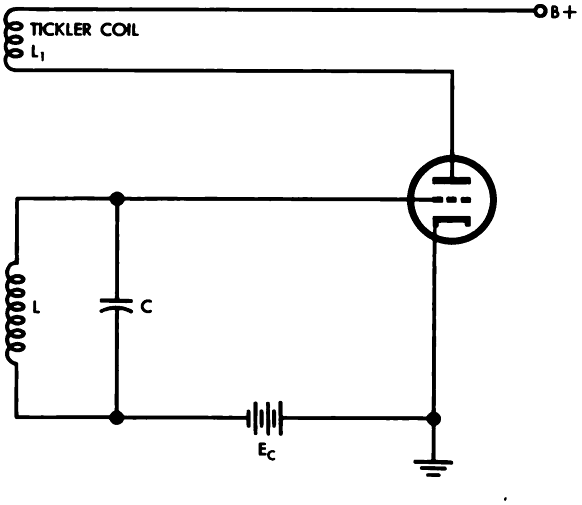

ARRANGED TO MAKE CONTACT ONCE EACH CYCLE AT POINT B
ENERGY SUPPLIED BY BATTERY BY USING A SWITCH
A
A'
A''
B
B'
B''

Producing Undamped Oscillations

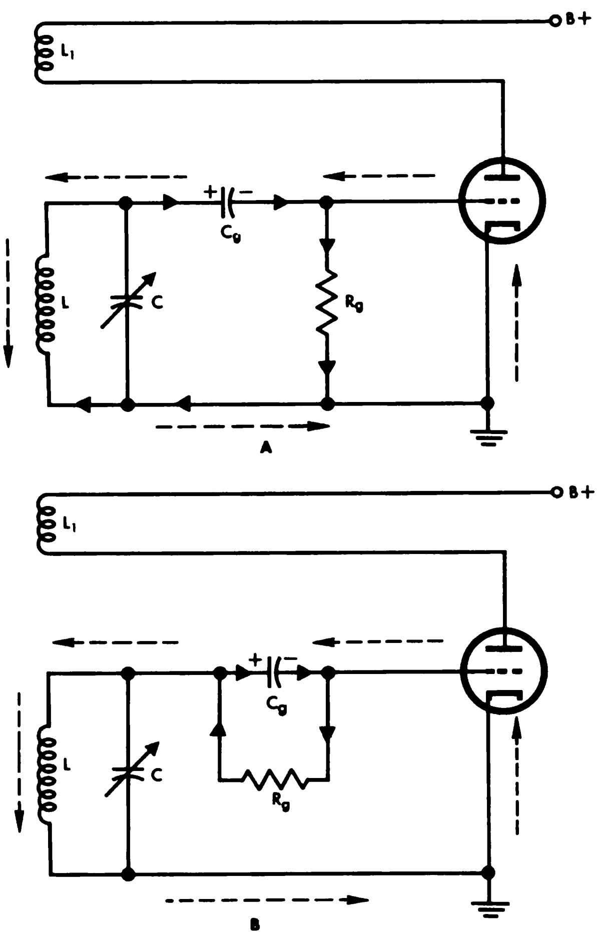
REPLACING THE SWITCH WITH A TUBE

An amplifier tube with its regenerative feedback path may be used in such a manner that the plate voltage will supply the energy dissipated by the resistance in the tank, thus replacing the switch and battery in the hypothetical circuit. This electronic switch has the advantage of applying the necessary pulses of energy at the proper instant to cause the tank circuit, through the *flywheel effect*, to deliver an output of sine-wave oscillations at a constant amplitude. In the illustration to the right, a tank circuit is connected to the grid of an amplifying tube with a coil in the plate circuit inductively coupled to the tank. This coil (L1), through which feedback occurs, is called a *tickler coil* and the circuit is known as a *tickler-coil oscillator*. The tickler coil connections are so made that the energy fed back to the grid tank from the plate is shifted 180°. Since the plate voltage is shifted approximately 180° from the grid voltage (through the amplifying tube), a total shift of 360° occurs in any signal applied to the grid. Therefore, the signal returned to the grid is an amplified signal in phase with the original signal. If the amount of regenerative feedback is sufficient, any oscillations set up are sustained.

Since no signal is applied to the grid, what causes oscillations to start? This question may be answered as follows: Electrons emitted from



A Simplified Tickler-coil Oscillator



Tickler Coil Oscillator with Grid Leak Bias

the cathode of a tube arrive at the plate at irregular intervals. Because of the irregularities in the flow of plate current, tube noise is generated. These noise currents contain all frequencies; therefore, voltages of all frequencies are induced in the grid tank coil. Those components of the induced voltage, which have the same frequency as that of the tank, supply energy to the tank. Through regenerative feedback, a signal is built up on the grid sufficient to set up oscillations. If the amplification of the tube is insufficient to supply the losses which occur in the circuit, the voltage across the tank coil will not build up and oscillations will not occur. However, if the feedback is of the proper phase and amplitude to overcome the losses in the circuit, oscillations will be started and sustained.

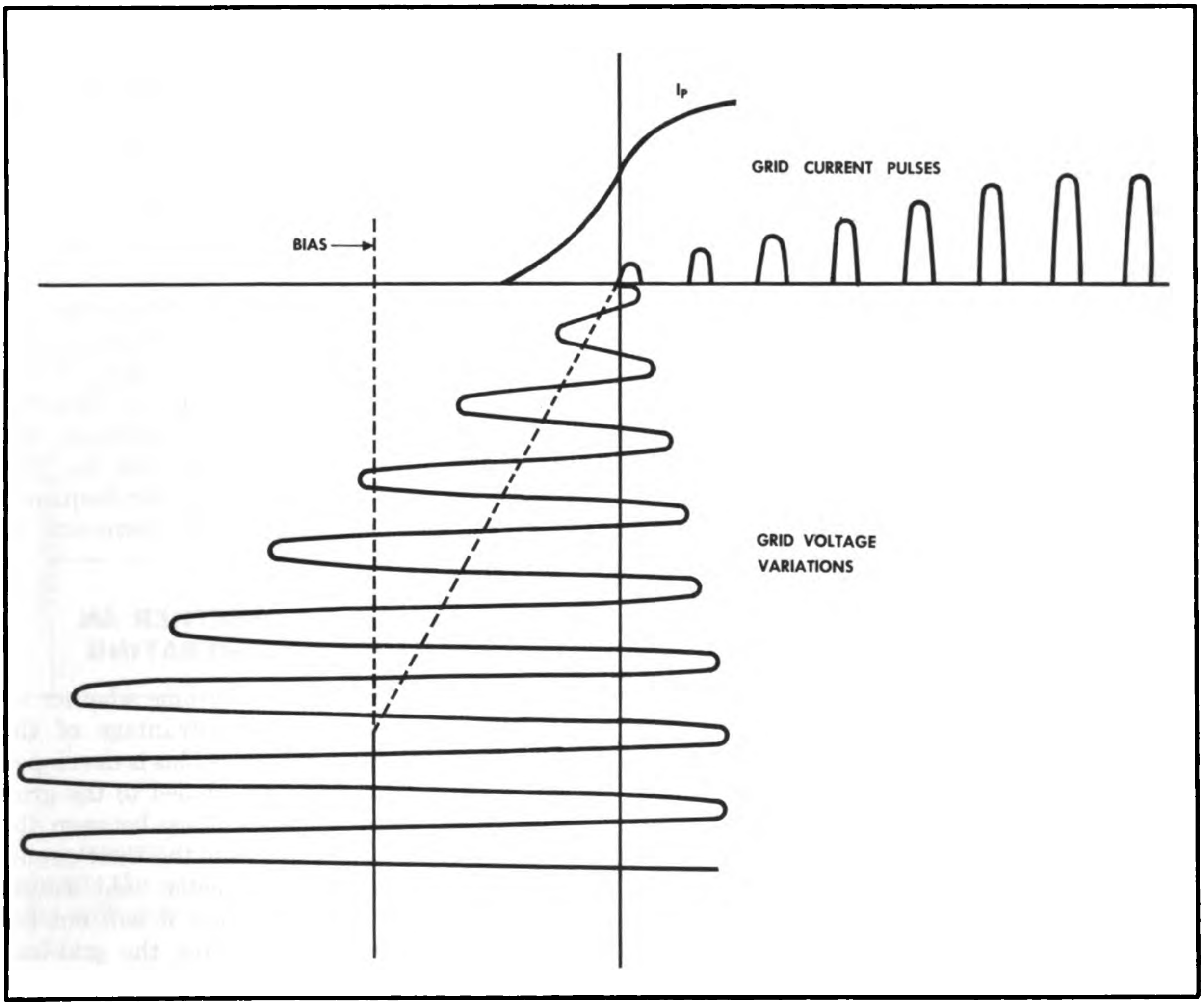
An oscillator in which oscillations are self-started and sustained is known as a *self-excited* oscillator. The basic requirements for any self-excited feedback oscillator must include:

1. An amplifier and its power supply.
2. Feedback in proper phase and amplitude.
3. An energy storage device to sustain oscillations and determine frequency.

GRID-LEAK BIAS IN OSCILLATORS

Although the tickler-coil oscillator was shown to be operating class A with fixed bias, its efficiency is not sufficient to generate appreciable power. Some laboratory oscillators (where a pure sine wave rather than much power is desired) are occasionally operated class A with fixed bias. However, most oscillators are operated class C in order to generate power with greatest efficiency. Class C operation was defined (on page 219) as the operation of a tube with a bias such that plate current flows for much less than 180° of a cycle. For class C operation, the tube is biased below cutoff. If you used fixed bias for class C operation, oscillations could not be self-starting. To overcome this difficulty, grid-leak bias is used in most oscillators. The illustration to the left shows the tickler-coil oscillator redrawn with grid-leak bias replacing the fixed bias. The grid-leak biasing system consists of grid capacitor C_g and grid-leak resistor R_g. Two methods of connecting the grid-leak resistor are shown in the illustration. Either may be used to achieve the same results.

Let us see how the grid-leak bias arrangement makes it possible for the oscillator to operate class C and still maintain its self-starting ability. When the circuit is first placed into operation, the grid-bias voltage is zero. The irregularities in the flow of current to the plate are coupled into the tank circuit and start oscillations. These oscillations are applied to the grid of the tube. As the grid goes positive with respect to the cathode, the grid draws current in pulses and the capacitor C_g charges, as shown by the dotted arrows. During the remainder of the sine wave, the charge on the capacitor leaks off (discharges) through the grid-leak resistor R_g. This discharge path is indicated by the solid arrows in the illustration. The capacitor-charging path has a time constant which is relatively short compared to the time constant of the discharge path. The capacitor therefore charges



Development of Grid Leak Bias

more during the time that grid current flows than it discharges during the remainder of the cycle.

You can see the buildup and attainment of proper bias in the above illustration. Note that the increase in bias is cumulative and continues until a constant bias is established. This occurs when the charge on the capacitor builds up to a value such that the addition of charge during each cycle exactly equals the charge that leaks off. The magnitude of the average bias may be controlled by varying the value of R_g. If the values of R_g and C_g are correctly chosen, the bias voltage builds up to, and remains around, an average value as long as a signal is being applied to the grid from the tank. This average value of bias keeps the tube biased much below cutoff, causing it to operate class C.

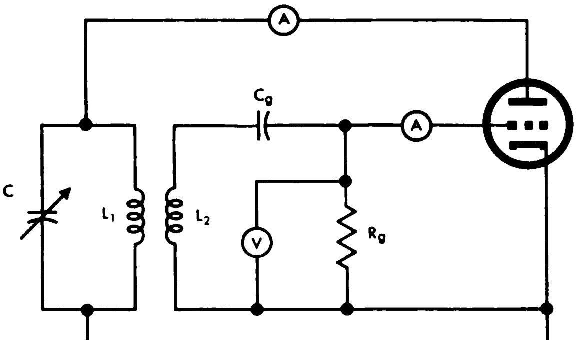
In addition to the increase in efficiency brought about by operating the oscillator class C, grid-leak bias helps to stabilize the amplitude of the oscillations in the tank. Since the bias follows any increase or decrease in the amplitude of the signal applied to the grid, the grid current becomes greater, and the bias increases if the signal applied to the grid tends to increase. This, in turn, causes the plate current and the amplitude of the oscillations to decrease. In similar fashion, a decrease in the amplitude of oscillation decreases the bias and tends to increase the amplitude of oscillations. Since the oscillator operates class C, plate current flows in pulses which occur once in each cycle. These pulses of current through the tickler coil induce pulses of voltage into the tank at the frequency at which oscillations are occurring and provide energy to sustain the oscillations of the tank. The result of proper grid-leak bias is a steady grid bias, equal amplitude plate current pulses, and oscillations of a constant amplitude.

Although the values chosen for R_g and C_g are not too critical and may be determined by experiment, the capacitor should offer negligible impedance compared to that of the resistor at the frequency of oscillation. However, the time constant R_gC_g must not be made too long or the charge accumulated on the condenser each time the grid draws current will not leak off during the remainder of the cycle. The bias then will keep building up on each cycle until it cuts the plate current off completely and oscillations cease. There will be no oscillations until the charge on the condenser leaks off sufficiently to allow oscillations to start up again. This intermittent operation results in an output which consists of a few cycles of oscillation followed by a period of no oscillation.

TUNED PLATE OSCILLATOR

The tickler coil oscillator is an example of the simplest type of RF oscillator. For radio-frequency operation, the values of L and C are made low by using an inductance with an air core and a capacitor with an air dielectric. A variable capacitor is used in most tank circuits so that the oscillator may be tuned to a desired frequency. When the tank of a tickler-coil oscillator is placed in the grid circuit, the oscillator is called a *tuned-grid* RF oscillator.

If the tuned tank is placed in the plate circuit instead of in the grid circuit, the oscillator is called a tuned-plate oscillator. Study the circuit of the tuned-plate oscillator shown on this page. Its operation is similar to that of the tuned-grid oscillator. Feedback is obtained through the mutual inductance between the tuned coil L_1 and the untuned coil L_2. The capacitor C is variable. It varies the frequency of oscillation by changing the frequency to which the tank is resonant.



A Tuned-plate Oscillator

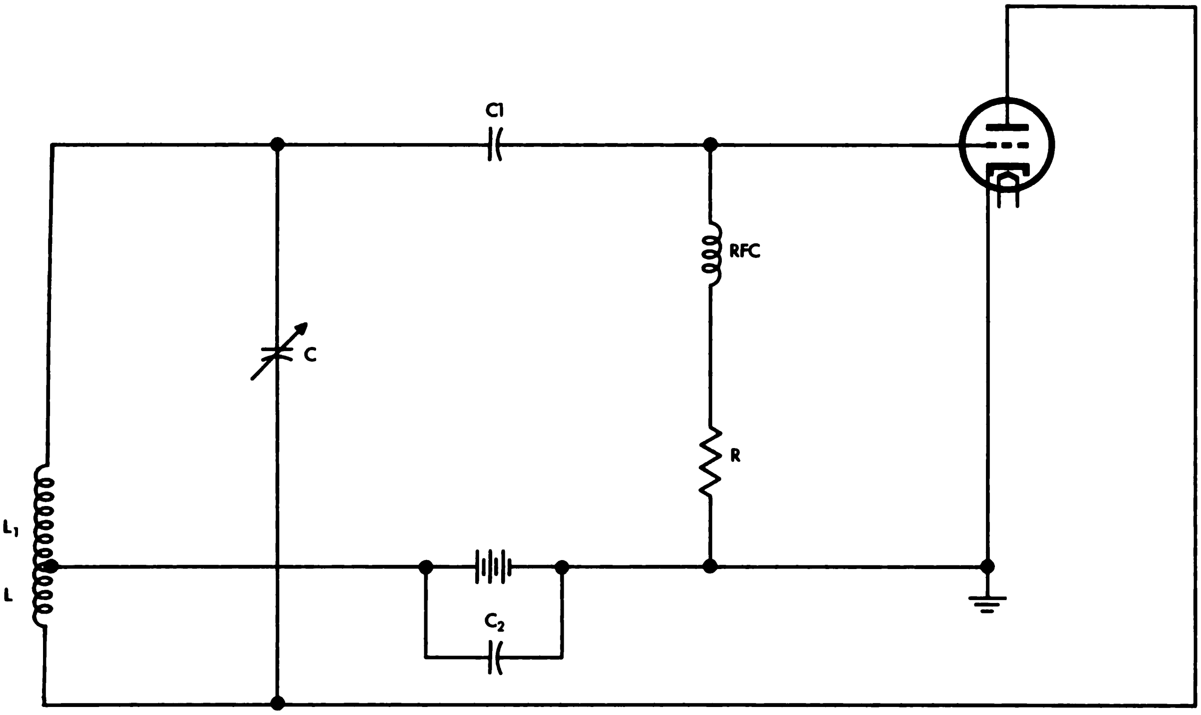
DETERMINING WHETHER AN OSCILLATOR IS OPERATING

Methods often used to determine whether an oscillator is operating take advantage of the grid-leak bias being used. Since bias is developed only when a signal is being applied to the grid, the presence of a negative voltage between the grid and ground indicates that the tank circuit is oscillating. The DC voltmeter used should have a high impedance, so that it will not act as a low-resistance shunt across the grid-leak resistor.

In some cases, a DC ammeter is used in the grid-circuit to indicate whether the oscillator is operating properly. The presence of grid current occurs only if the grid is being driven positive; therefore, grid current indicates an oscillating condition.

A third method used to indicate oscillating conditions employs a DC ammeter in the plate circuit. Since the grid-leak bias causes the tube to conduct in pulses, the average plate current becomes much less when oscillations are present than when there are no oscillations and the grid bias is zero. A dip in the plate current reading, therefore, indicates oscillating conditions.

Note how the meters are connected in the circuit of the tuned-plate oscillator. Any one of these meters will indicate an oscillating condition. A gas-filled tube, such as a neon tube, may also be used to detect the presence of RF oscillations. If such a tube is brought close



A Hartley Oscillator

to a resonant tank in which oscillations are present, the gas in the tube will ionize and the tube will glow.

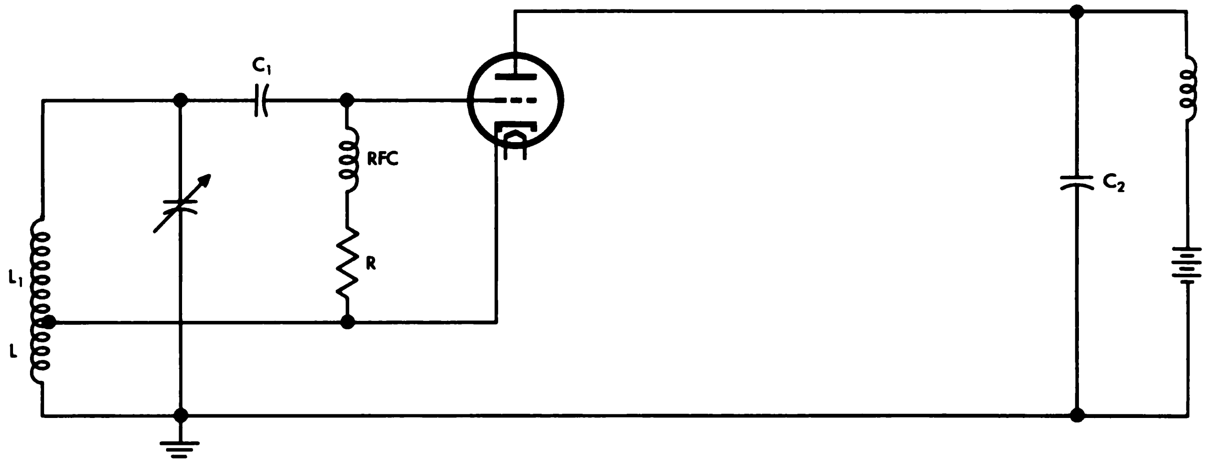
HARTLEY OSCILLATOR

As you have seen, any amplifier circuit in which feedback of proper phase and amplitude takes place and which contains an energy storage device may be used to generate sine waves. The tuned-plate and tuned-grid oscillators are examples which serve to illustrate these basic principles. However, modifications may be made in oscillator circuits which result in more efficient operation. Although all feedback oscillators operate on basically the same principles, many different types are found, each of which has certain advantages over the others.

The Hartley oscillator is a popular variable frequency oscillator, simple in design and very dependable. It gives a fairly uniform output for a range of frequencies from about 800 kc to 100 mc. In addition to its widespread use in transmitters, it is in common use in television and FM receivers. In the schematic diagram of the Hartley oscillator shown above, note that there is a single coil instead of two as in the tuned grid oscillator. One end of the tank coil is connected to the plate, the other end of the coil is connected to the grid, and the cathode is connected near the plate end. The tank capacitor shunts both sections of the coil (L and L1). Feedback is accomplished through inductive coupling between L and L1.

When you examine this circuit, you will see that it is similar to that of the tuned grid oscillator discussed previously. The only difference is in the arrangement of the tickler coil and the tuning capacitor. In the Hartley, the single coil with a tap near the cathode end takes the place of the two separate coils used in the tuned grid oscillator. The feedback can be controlled by adjusting the tap on the coil to vary the number of turns between the plate and the cathode. Note, too, that the entire coil is tuned by capacitor C and together they form the resonant circuit. The tank circuit is thus common to both plate and grid circuits.

The initial starting action is similar to that of the tuned grid oscillator. Coil L, L1 functions as an autotransformer. When the plate voltage



Modified Hartley Oscillator

is applied, a surge of plate current flows through L. An alternating current is caused to flow in the tuned circuit by the transformer action. The oscillating current flowing in the tuned circuit produces a voltage drop across L and L1. Because the lower end of L1 is connected to the cathode through bypass capacitor C2, the AC voltage across L1 is applied to the grid-cathode circuit. This AC voltage causes the plate current to vary at the same frequency as the alternating current within the tank.

The RF choke has high impedance to RF current and low resistance to DC. Therefore, it permits only the DC component of the rectified grid current pulses to flow though it and the grid resistor R. Bias builds up in the same way as in the tuned grid oscillator, and soon the tube is biased for class C operation.

Modified Hartley Oscillator

Another type of Hartley circuit is the modified Hartley. Notice that the modified Hartley oscillator is arranged so that one side of the tuning capacitor is grounded. The plate of the tube is maintained at RF ground potential by the bypass capacitor C2, and the RF component of the plate current flows from cathode to plate, through the bypass capacitor to L, and back to the cathode. With this arrangement, a ganged capacitor may be used for tuning, and the rotors of the ganged sections may be common to ground. Keeping the rotor of the capacitor at ground potential in this way eliminates the harmful effects of hand capacitance. If the rotor is not grounded, you add capacitance to the resonant circuit when your hand is at the rotor tuning it. Then, when you remove your hand, you remove the hand capacitance, and the circuit is no longer tuned to the desired frequency.

One disadvantage of the modified Hartley is that the cathode is not at RF ground potential. The heater is usually at ground potential, and the capacitance between the cathode and the heater is in parallel with a portion of the tank coil. As the temperature of the heater changes, the interelectrode capacitance is altered and the frequency of the oscillator is changed. To avoid this trouble when using a modified Hartley in transmitters requiring high-frequency stability, special care must be taken to maintain the heater at cathode RF potential by placing RF chokes in the heater leads.

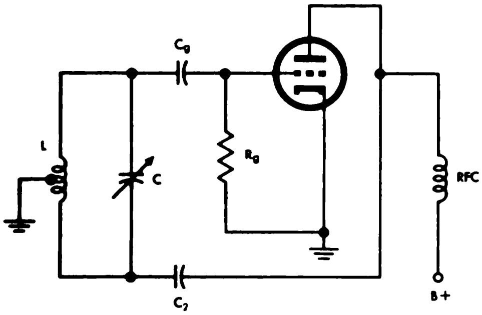
The tank circuit of the Hartley oscillator is tuned by means of the capacitor. A tank circuit may normally be tuned by varying either the inductance or the capacitance, but in the Hartley oscillator only capacitance tuning is used. Varying the inductance would be inconvenient, because a change in tuning would cause a change in grid excitation unless the inductance of both grid and plate coils were variable.

Note that the B+ power supply is connected so that it is in series with the plate-to-cathode circuit of the tube and a portion of the tapped RF coil in both types of Hartley oscillators. When connected in this way, a Hartley oscillator is said to be *series-fed*. Note also that both the DC and RF currents flow through part of the RF coil. This part of the tank coil (L) is therefore at a high DC potential with respect to ground, and the tuning capacitor must be able to withstand the sum of both the DC and RF voltages. Because of this disadvantage, series-fed Hartley oscillators are used chiefly in low-power circuits where low voltages are used. The capacitor C_2 is used to by-pass the large RF currents around the B+ power supply.

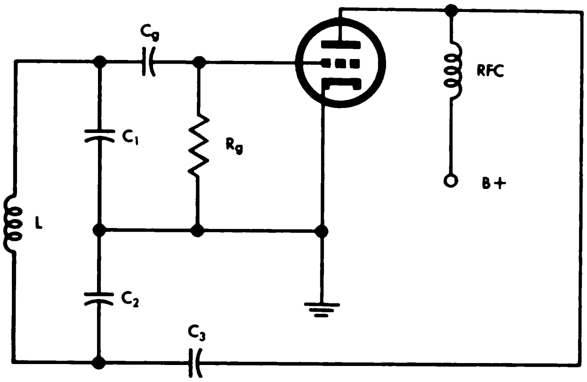
Shunt-Fed Hartley Oscillator

To overcome the main disadvantage of the series-fed Hartley oscillator, a shunt-fed Hartley oscillator may be used. As you can see in the illustration below, the B+ plate supply is connected in parallel with the tank circuit between plate and cathode. An RF choke coil is placed in series with B+ to prevent RF variations in the power supply. It also prevents the power supply from short-circuiting the plate to the cathode at RF potential. A bypass capacitor C_2 is used to prevent the tank coil from short-circuiting the power supply. The capacitor is made large enough to offer negligible impedance to the RF current, and it must have a voltage rating which is sufficient to withstand the sum of the DC and RF plate potentials.

Note that in the shunt-fed Hartley circuit the RF and DC current paths are different and are in parallel. The RF current flows from the plate of the tube through the by-pass capacitor, through the lower portion of the tank to ground. From there it continues to the cathode, through the tube, and then back to the plate.



Shunt-fed Hartley Oscillator



A Colpitts Oscillator

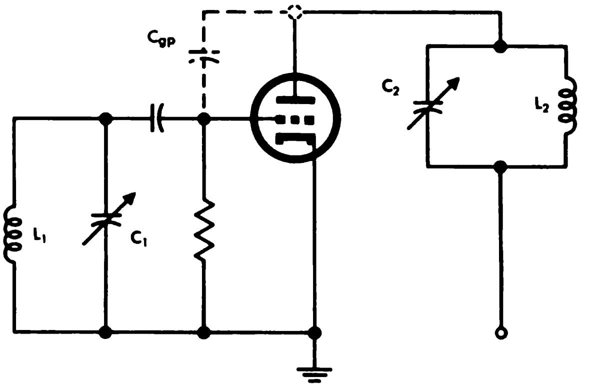
The DC current flows from the plate of the tube through the RF choke and through the power supply to ground. From there it continues, like the RF current, to the cathode, through the tube, and back to the plate. Observe that only RF current flows through the tank, and the tank circuit is therefore at a low DC potential. In this way, damage to the circuit and injury to operating personnel, which may occur in the series-fed Hartley, is prevented. The shunt-fed Hartley may therefore be used in high-power oscillators.

THE COLPITTS OSCILLATOR

Another type of oscillator commonly found in radio and radar circuits is known as a Colpitts oscillator. It is very much like the Hartley in construction except that the capacitive branch of the tank circuit is tapped. As you can see in the illustration above, the cathode is connected between the two capacitors C_1 and C_2 which are connected in series.

Note that grid-leak bias is used in this circuit, too. As in the shunt-fed Hartley oscillator, an RF choke and a by-pass capacitor (C_3) are used to keep oscillations out of the DC power supply. A Colpitts oscillator must be shunt-fed, because there is no path for DC through the tank back to the tube. DC current flows from the plate directly back through the tube in shunt with the RF current path. Plate voltage and grid voltage are 180° out of phase; thus, the feedback is in proper phase to sustain oscillation. The gain of the tube must be sufficient to overcome losses in the circuit and depends upon the ratio of C_1:C_2.

The frequency at which oscillations occur is the resonant frequency of the tank and is equal



A Tuned-plate, Tuned-grid Oscillator

to approximately $\frac{1}{2\pi\sqrt{LC}}$, where the value of the capacitance C is determined by C_1 and C_2 in series, $\frac{C_1 C_2}{C_1+C_2}$. The circuit is, therefore, a little more critical to tune than the Hartley if variable capacitors are used. This is due to the fact that each time C_1 is varied, C_2 must also be adjusted to keep the feedback constant. To offset this disadvantage, the capacitors C_1 and C_2 are usually ganged. In most cases, inductive tuning is used and the capacitances are fixed at the desired ratio.

Good frequency stability is obtained in the Colpitts oscillator by making the tank capacitances C_1 and C_2 of comparatively large value. The tube and circuit capacitances which are in parallel with the tank capacitances and comparatively low in value will therefore not affect the total capacitance to any great extent.

TUNED-PLATE TUNED-GRID OSCILLATOR

Another type of oscillator found in electronic equipment employs two separate tank circuits. The tuned-plate, tuned-grid oscillator shown above is one such type. For convenience in referring to it, the abbreviation TPTG is often used. At first glance, there appears to be no feedback path from plate to grid, since L_1 and L_2 are not inductively coupled. However, feedback from plate to grid occurs through the plate-to-grid interelectrode capacitance (C_{gp}), shown as a broken line in the illustration.

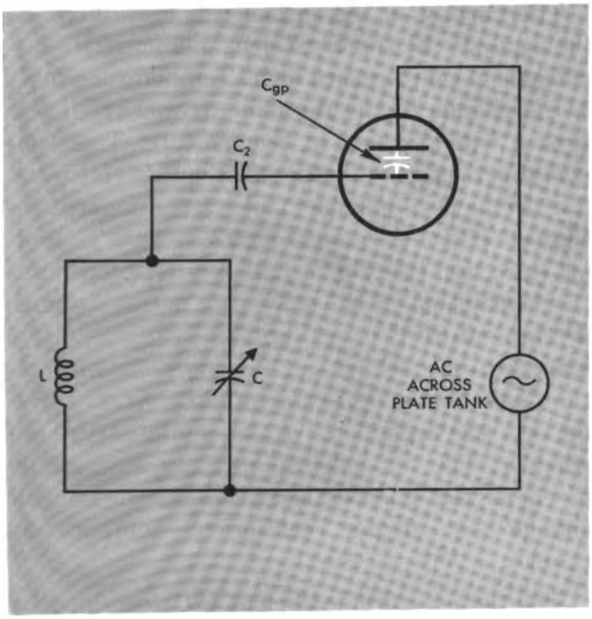
When the plate voltage is applied, the variations in plate current set up oscillations in the plate tank. These oscillations cause an RF voltage to be developed across the plate tank. This changing RF voltage may be represented as an RF generator supplying energy to the grid tank. As you can see in the illustration on page 245, this generator causes current to flow through C_{gp} and through C_2 to the tuned-grid circuit. Since grid-leak bias is used in the same way as in the other oscillators, energy is supplied to the plate tank in pulses.

For the feedback from plate to grid to be in proper phase to sustain oscillations, both tanks must be tuned somewhat inductively. This means that the plate and grid tanks are tuned to a frequency which is a little higher than the output frequency. If the resonant frequency of either tank is changed, the frequency of oscillation will change slightly but remain lower than the resonant frequency of both tanks.

Since the frequency of oscillation is only slightly lower than that of the tank with the lower resonant frequency, you can obtain a good approximation of the output frequency by computing the resonant frequency of this tank. This may be done by using the formula,

$$f = \frac{1}{2\pi\sqrt{LC}}$$

Usually a DC plate current meter is used to indicate proper tuning of a TPTG oscillator. The graph to the right on page 245 shows how the plate current varies as the tuned plate circuit is adjusted. With the capacitor set at minimum capacitance, the circuit is inductive and oscillations occur if the tank circuit impedance is great enough. As the plate tuning capacity is gradually increased toward point B and resonance is approached, the impedance increases and so does the excitation voltage. Increasing the excitation voltage increases the bias. This causes the DC plate current to decrease. Thus, when the capacity is increased to point B — the point where the circuit is tuned to resonance — the plate current is minimum and the oscillations are at their greatest magnitude. If the capacity is increased beyond point B (toward point D), the plate current increases and the magnitude of oscillation decreases. At point D, the circuit is no longer inductive and oscillations cease. When this method of tuning is used, the plate is not tuned for minimum, but to a point indicated by A or C on the graph. Though the DC plate current here is slightly higher



Energy Supplied to Grid Tank of TPTG Oscillator

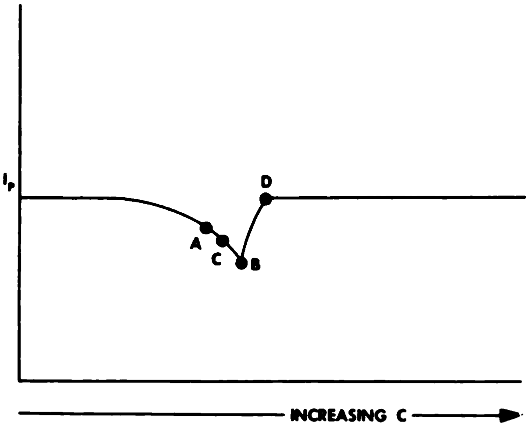
than minimum, the oscillator is more stable in operation at this point. If the circuit were tuned to minimum plate current (point B), any change might alter the reactance of the circuit and cause the oscillations to stop.

CRYSTAL OSCILLATORS

When the frequency of a transmitter must be held within close limits, a crystal-controlled oscillator is used. This is an oscillator in which a specially cut crystal controls the frequency. Crystal-controlled oscillators are the standard means used for maintaining the frequency of radio transmitting stations within assigned limits. For the accuracy in timing needed in radar, a crystal-controlled oscillator is usually employed as a basic circuit to generate an output at a highly stable, accurately measured frequency. Crystals are used also as the active element in high-fidelity microphones, loudspeakers, and electric pick-up arms.

Piezoelectric Effect

Crystals exhibit a characteristic which makes possible their various uses. This characteristic, known as *piezoelectricity* or the *piezoelectric effect,* is the property of a crystal by which mechanical stresses produce electric charges, and conversely, electric charges produce mechanical stresses. If a voltage is applied to a crystal, it will vibrate at the frequency of the applied voltage. The amplitude of vibration of a given crystal, however, is much larger at one frequency than at all others. This frequency is the *resonant* frequency of the crystal.



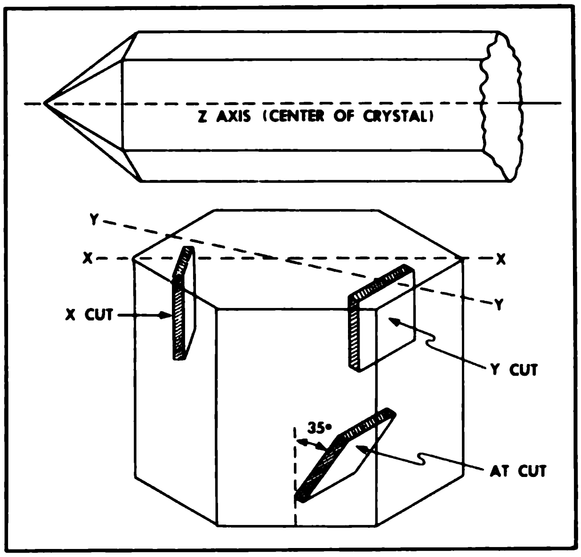
Effect of Increasing Capacitance on Plate Current

This piezoelectric effect is exhibited to different degrees by crystals of various substances such as quartz, tourmaline, and Rochelle salt. Of the three, Rochelle salt exhibits this effect to the greatest degree. Quartz crystals, however, are used for frequency control in most oscillators, because quartz has greater mechanical strength.

A natural quartz crystal has a hexagonal cross section and pointed ends, as shown, page 246. Sections cut from such a crystal exhibit the property of piezoelectricity, which may be expressed in terms of three sets of axes. The axis joining the points at the ends of the crystal is known as the *optical* or *Z-axis.* Mechanical stresses along this axis produce no piezoelectric effect. A hexagonal cross-section results when a quartz crystal is cut perpendicular to its optical axis. The three axes passing through the corners of this hexagon are called *electrical* or *X-axes,* while the three axes perpendicular to the faces of the crystal are known as *mechanical* or *Y-axes.* One X-axis and one Y-axis are shown in the illustration on the next page.

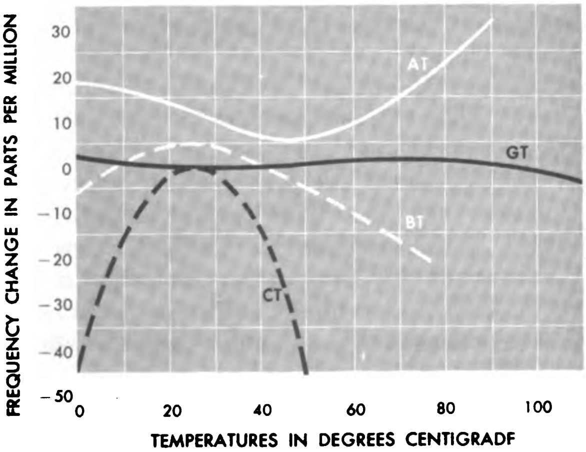
Crystal Cuts

The crystals used in electrical circuits are thin sheets cut from the natural crystal and



Quartz Crystal Cuts

are ground to the proper thickness for the desired resonant frequency. For any given crystal cut, the thinner the crystal, the higher is the resonant frequency. The cut of the crystal means the definite way in which the usable crystal is cut from the natural crystal. Crystals are cut at various angles with respect to the X, Y, or Z-axis. The various types of cuts are X, Y, AT, BT, CT, DT, and GT. Only the X, Y, and AT cuts are shown in the illustration. In the X and Y cuts, the face of the crystal is cut parallel to the Z-axis. AT, BT, CT, DT, and GT are called special cuts. They are cut with the face of the crystal at an angle to the Z-axis.



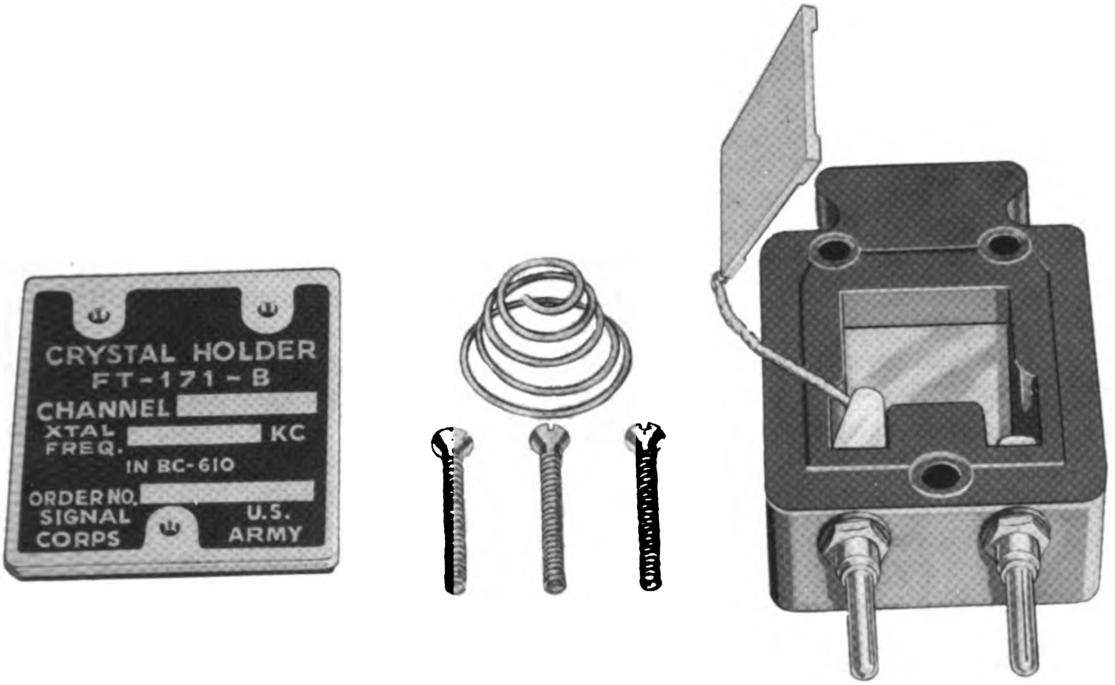
Effect of Temperature on Frequency for Typical Crystal Cuts

The purpose of the special cuts is to increase the stability of the crystal by reducing its coefficient of temperature drift. Temperature drift is the change in the resonant frequency of the crystal with changes in temperature. The *temperature coefficient* is the relationship between the frequency change and the temperature change. If the frequency increases with a temperature rise, the crystal has a positive temperature coefficient. If the frequency decreases with a temperature rise, the crystal has a negative temperature coefficient. If variations in temperature have little or no effect on resonant frequency, the crystal has a zero temperature coefficient. The temperature coefficient is usually given as the number of cycles per megacycle change per degree centigrade temperature change.

The temperature coefficient is determined by the cut. The X-cut has a negative temperature coefficient. The Y-cut has a positive temperature coefficient. The AT-cut has a low or practically zero temperature coefficient. The GT-cut has a zero temperature coefficient over a wide temperature range. The effect of variations in temperature on the frequency of some typical crystal cuts may be seen in the illustration at the lower left.

Transmitters which require a very high degree of frequency stability, such as broadcast transmitters, use temperature-controlled ovens to maintain a constant crystal temperature. These ovens are thermostatically controlled containers in which the crystals are placed. Of course, crystals with special cuts are used with a temperature coefficient as close to zero as possible.

The type of cut also determines the activity of the crystal. Some crystals vibrate at more than one frequency and thus will operate at harmonic frequencies. These crystals are commonly called crystals with multiple peaks. Crystals which are not of uniform thickness may have two or more resonant frequencies. Usually one resonant frequency is more pronounced than the others, and the others are called spurious frequencies. Sometimes such a crystal oscillates at two frequencies at the same time. In other cases, the crystal jumps frequency as the plate tank capacitor is varied past a certain point.



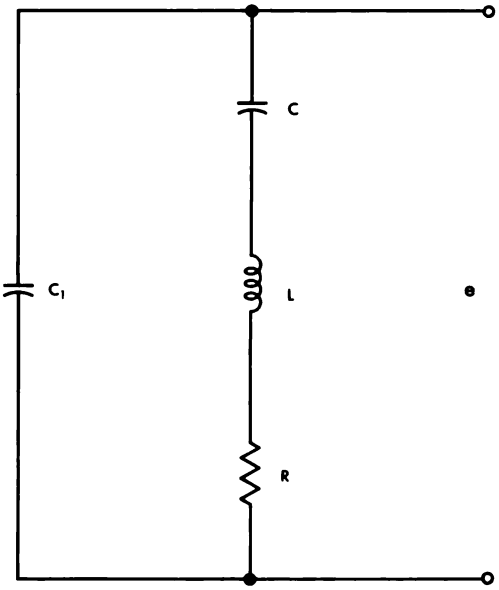
Typical Crystal Holder

The amount of current that can safely pass through a crystal ranges from 50 to 200 milliamperes. When the rated current is exceeded, the amplitude of mechanical vibration becomes too great and the crystal may crack. Overloading the crystal affects the frequency of vibration, because the power dissipation and crystal temperature increase with the amount of load current.

Crystals In Circuits

Crystals are mounted in holders which support them physically and provide the electrodes by which voltage is applied to them. In addition, the holder must allow the crystal maximum freedom of vibration at all times. Many types of crystal holders are used for this purpose. The one shown in the illustration above is one type of holder used in mounting crystals. Note that in this holder the crystal is placed between two parallel plates with a small air gap between the crystal and the top electrode. This arrangement may be used with AT and BT crystal cuts as well as in cases where the crystal is not subjected to mechanical shock. Crystal holders have a significant effect on the frequency. In fact, some variable gap holders shift the frequency of a 14-mc crystal as much as 20 to 50 kc when the air gap is varied.

The crystal and its mounting are usually enclosed in a sealed container for protection from moisture and other climatic hazards. In some cases, the holder is evacuated and provided with a heater and thermostat (electric oven) to maintain the temperature of the crystal constant.



Equivalent Circuit of Quartz Crystal and Holder

The frequency stability of crystal-controlled oscillators depends on the Q of the crystal and its temperature coefficient. The Q of a crystal is very high. It may be more than 100 times greater than that obtained with an equivalent electrical circuit. The Q of a crystal is determined by the cut, the type of holder, and the accuracy of grinding. Commercially produced crystals range in Q from 6000 to 30,000, while some laboratory experimental crystals range in Q up to 400,000.

The following data show the relative size and characteristics of a typical crystal.

Dimensions
- Thickness — .25 inch
- Width — 1.3 inches
- Length — 1.08 inches

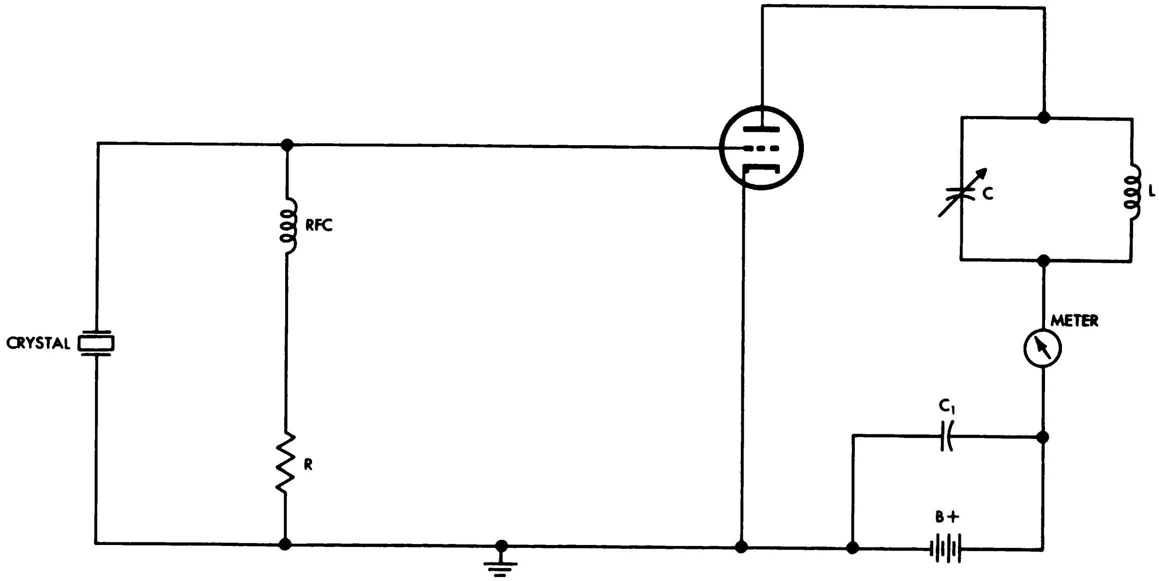
Resonant frequency — 430 kc

Equivalent electrical characteristics
- Inductance — 3.3 h
- Capacitance — 0.042 mmf
- Capacitance of holder — 5.8 mmf
- Q — 23,000

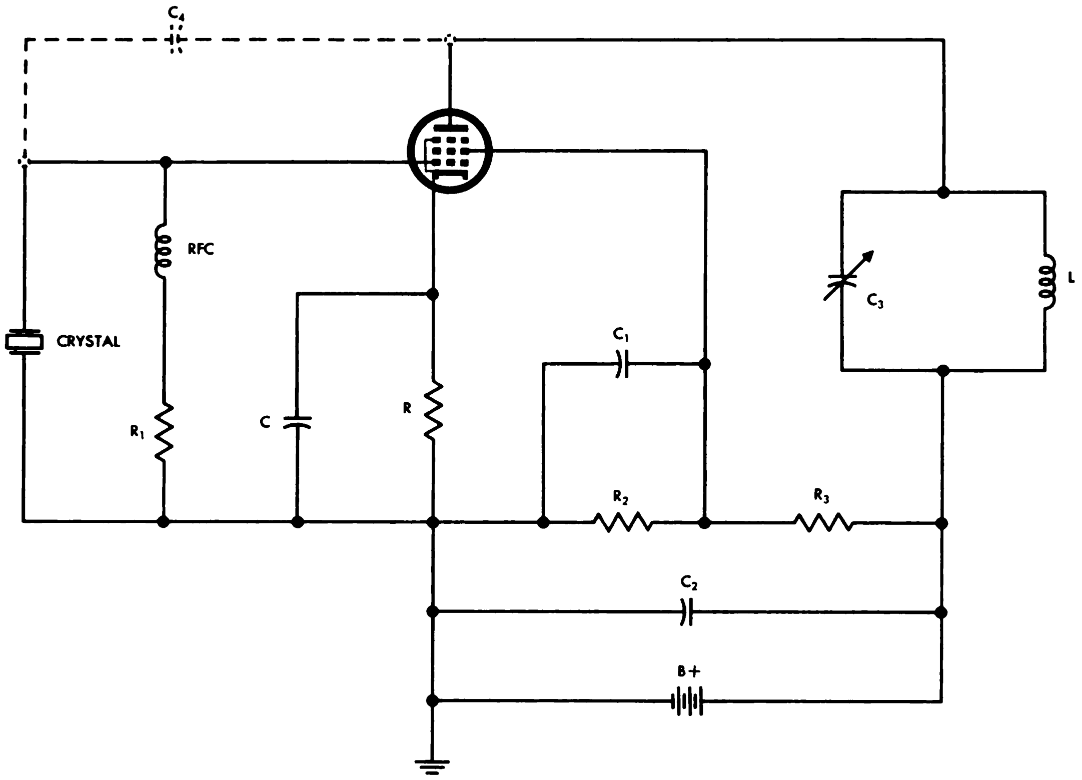
When placed in an electrical circuit, a crystal can act like a very high Q series resonant circuit. The electrical circuits associated with a vibrating crystal can be represented by an equivalent circuit composed of resistance, inductance, and capacitance, as shown at the bottom of the previous page. In this circuit, C_1 represents the capacitance between the metal plates of the crystal holder. When the crystal is not vibrating, the circuit acts only as this capacitance. The series combination of L, C, and R represents the electrical equivalent of the vibrating crystal characteristics. The frequency stamped on a crystal holder is usually the parallel resonant frequency of the crystal and its holder as applied in an oscillator circuit. This does not mean that an ordinary coil and capacitor could be substituted for the crystal to give the same electrical characteristics. The Q of the crystal circuit is many times higher than that of an equivalent LC-type circuit. This high Q factor gives circuit stability to the crystal which could never be obtained in an ordinary tank circuit.

Crystal-Controlled Oscillator

Since a vibrating quartz crystal is equivalent to a resonant circuit, it can be used in place of the usual tuned circuit as a frequency controlling element in an electron tube oscillator. The circuit shown below represents the common type crystal oscillator. It is similar to the tuned-plate, tuned grid-oscillator, with the crystal replacing the tuned-grid circuit. The feed-back takes place through the plate-to-grid interelectrode capacitance. The oscillations occur at the resonant frequency of the crystal, and the plate circuit is tuned slightly higher, or inductively.



Crystal Controlled Oscillator



Pentode Crystal Oscillator

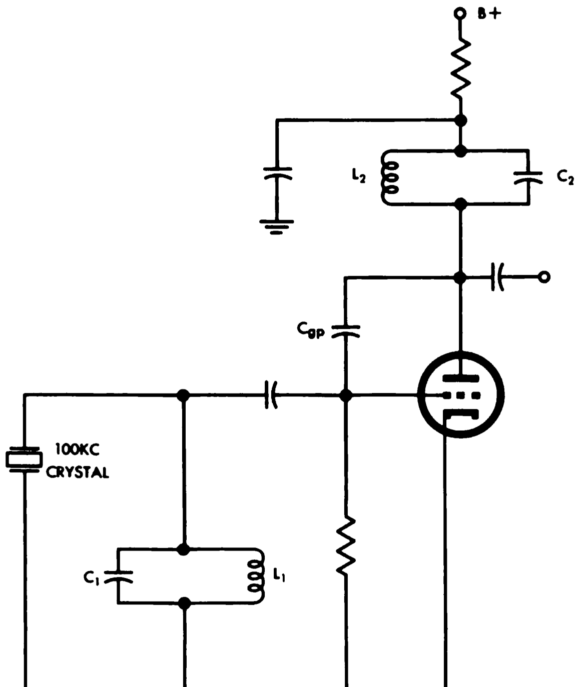
The operation of this circuit is similar to that of the tuned-plate, tuned-grid oscillator. When the filament has heated the cathode to the emission point, the plate voltage is applied, and the tube starts conducting. The surge of plate current flowing through L releases energy to the resonant plate circuit and starts it oscillating. The RF voltage developed by the alternating current in the plate tank is applied to the crystal through the plate-to-grid capacitance and bypass capacitor C_1. Placing this voltage across the two plates of the crystal holder causes the crystal to bend physically. When the first surge of plate current is over, the crystal springs back to its normal shape. In doing this, it generates a voltage of its own, causing the grid to become positive. The grid then draws current, becomes negative, and the cycle is repeated. If enough energy is applied back to the circuit in the proper phase to make up for resistive losses, the process continues at the mechanically resonant frequency of the crystal.

A resistor is connected from grid to ground through a choke to allow electrons which have been attracted to the grid on the positive swing to return to the cathode ground. The current flowing through this resistor develops the bias voltage. The capacitor usually associated with the grid leak resistor is not necessary, because the crystal and holder serve in its place.

In the tuning of a crystal oscillator, a meter is generally used. Note the crystal oscillator circuit. The meter is placed in series with the plate circuit. DC plate current flows from the cathode to the plate, then through the meter. Any change in plate current is thus shown on the meter (a DC milliammeter). When the oscillator is generating its greatest amount of RF current, the grid draws its greatest current, and the bias is maximum. This holds the plate current to a minimum, and the meter reading is at a minimum. Now you might think that the oscillator should be tuned for this minimum, for it is here that the amplitude of oscillations is greatest. In actual tuning, though, the plate is first tuned for minimum by adjusting the tuning capacitor in the plate tank; then it is detuned (the capacity is reduced) to cause about 3 to 5 ma more plate current to flow than at minimum.

The reason for not tuning the plate circuit to the exact crystal frequency is that exact tuning would result in erratic and unstable operation, as in the case of the tuned-grid, tuned plate oscillator. The plate circuit must be tuned far enough on the inductive side of resonance to prevent changes, in the load placed on the tank of the oscillator by secondary circuits of the transmitter, from stopping oscillation.

The power output that can be obtained from a crystal oscillator is limited. If too high an excitation voltage is applied to an active crystal, it may vibrate strenuously enough to shatter itself. For this reason, the triode tube crystal oscillator shown is seldom used. Usually, a pentode or tetrode is used instead, for pentodes and tetrodes are tubes of high power sensitivity. With them, a comparatively high power output may be obtained from a small excitation voltage. This means that only a small amount of excitation voltage is needed to drive the tube to full power output, and a relatively small amount of feedback is enough to meet the requirements for oscillation. The small grid-to-plate capacitance in these tubes offers high impedance to RF and limits the grid excitation to the small required value. In fact, this feedback by the interelectrode capacitance may be too small, and it may be necessary to add an additional capacitance between grid and plate.



AN/APN-9 Crystal Oscillator

The operation of the typical pentode crystal oscillator circuit shown on page 249 is identical to that of the triode just explained. Capacitor C4 represents a capacitor that may have to be added to supply sufficient feedback to keep the circuit oscillating.

Note that grid leak bias is supplemented by cathode bias in this circuit. Cathode bias is developed by resistor R. The DC flowing from cathode to plate develops a voltage across the resistor that makes the end connected to the cathode positive with respect to ground. Since the lower end of the cathode resistor is effectively connected to the grid by R1 and the choke, any voltage developed across R appears between the grid and the cathode. This makes the grid negative, with respect to the cathode, by an amount equal to the voltage drop across R. Capacitor C is a bypass capacitor to bypass the RF component of plate current around the cathode resistor. The cathode resistor R is sometimes referred to as a minimum bias resistor, for bias voltage is developed across it any time plate current flows. Thus it is independent of oscillation and prevents excessive tube current if oscillation stops.

Application in Observer Equipment

Crystal-controlled oscillators are used in radar equipment where accurate timing is needed. The output of the crystal-contolled oscillator is highly stable at an accurately measured frequency. The use of special frequency divider circuits makes it possible to count down the frequency of this output to produce outputs at many desired frequencies. These outputs may be used for measuring time in radar circuits.

The 10-microsecond markers which you have seen on the indicator of the AN/APN-9 are produced by a crystal-controlled oscillator whose output is stabilized at 100 kc. A simplified schematic of this oscillator is shown to the left. The output of this oscillator is counted down and used also to stabilize the 50-microsecond, 100-microsecond, 500-microsecond, 1000-micro-second, and 5000-microsecond markers.

Chapter 28

Principles of Transmission

In chapter 24, you were introduced in a general way to radio transmission and reception. You were told that a radio communications system consists of a transmitter, an antenna system, and a receiver. In this and the following chapter, you will be shown how the various electronic circuits you have studied (rectifiers, amplifiers, and oscillators) may be combined to perform the functions of transmission and reception.

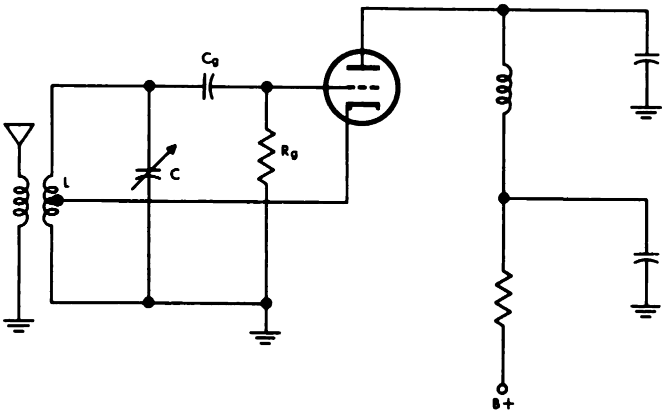
AN OSCILLATOR AS A TRANSMITTER

As you recall from chapter 24, a transmitter consists of two sections: a *radio-frequency section* to generate RF power, and a *modulator section* to vary the radio frequency energy in accordance with the information to be sent. Any oscillator (Hartley, Colpitts, crystal, etc.) may be used to produce the RF energy necessary to transmit a signal.

Producing a Carrier Wave

The schematic shown on the next page is a Hartley oscillator to which an antenna circuit has been added. The oscillations that are set up in the tank are coupled to the antenna circuit through the transformer. The alternating currents flowing back and forth in the antenna (which is a conductor) cause magnetic and electric fields to build up and collapse, then build up in the opposite direction and collapse again. Some of the energy in these fields returns to the conductor; however, most of the energy is radiated from the antenna in the form of electromagnetic waves.

Electromagnetic waves consist of electric and magnetic fields perpendicular to each other and to the direction in which the waves are traveling. The strength of these fields varies in the same manner and at the same frequency as the oscillations in the tank. They can there-



Producing a Carrier Wave

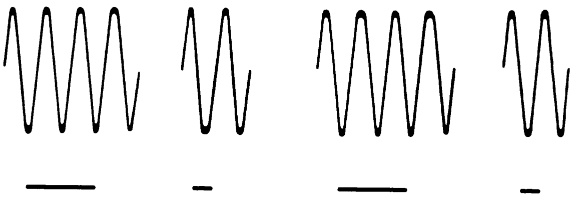
fore be represented by sine waves as shown in the illustration below. This wave is known as a *carrier wave* because it is used to carry intelligence from the transmitter to the receiver. Note that the carrier wave shown in the illustration does not vary in amplitude (strength) or frequency.

Methods of Modulating the Carrier Wave

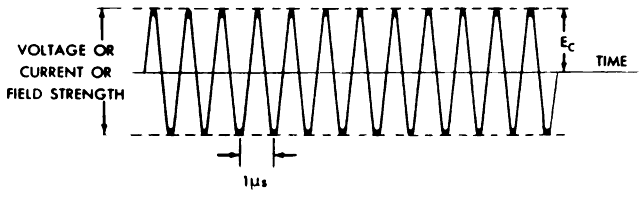
The carrier wave shown cannot be used to transmit information because its signal is unvarying. To send information, it is necessary to change or *modulate* the carrier wave in some way. There are several ways in which the carrier wave may be changed to transmit information.

One method, known as *keying*, consists of starting and stopping the continuous carrier wave so as to break it up into dots and dashes as shown in the illustration at the right.

The oscillator shown above may be used to send information by adding a device, or key, for interrupting the continuous carrier wave. By making and breaking the circuit at some point, oscillations may be started or stopped according to a code. The key may be placed in the plate, control grid, or power supply circuit. If a pentode is used, the key may be placed in the screen grid circuit. Coded messages may be sent in this way.



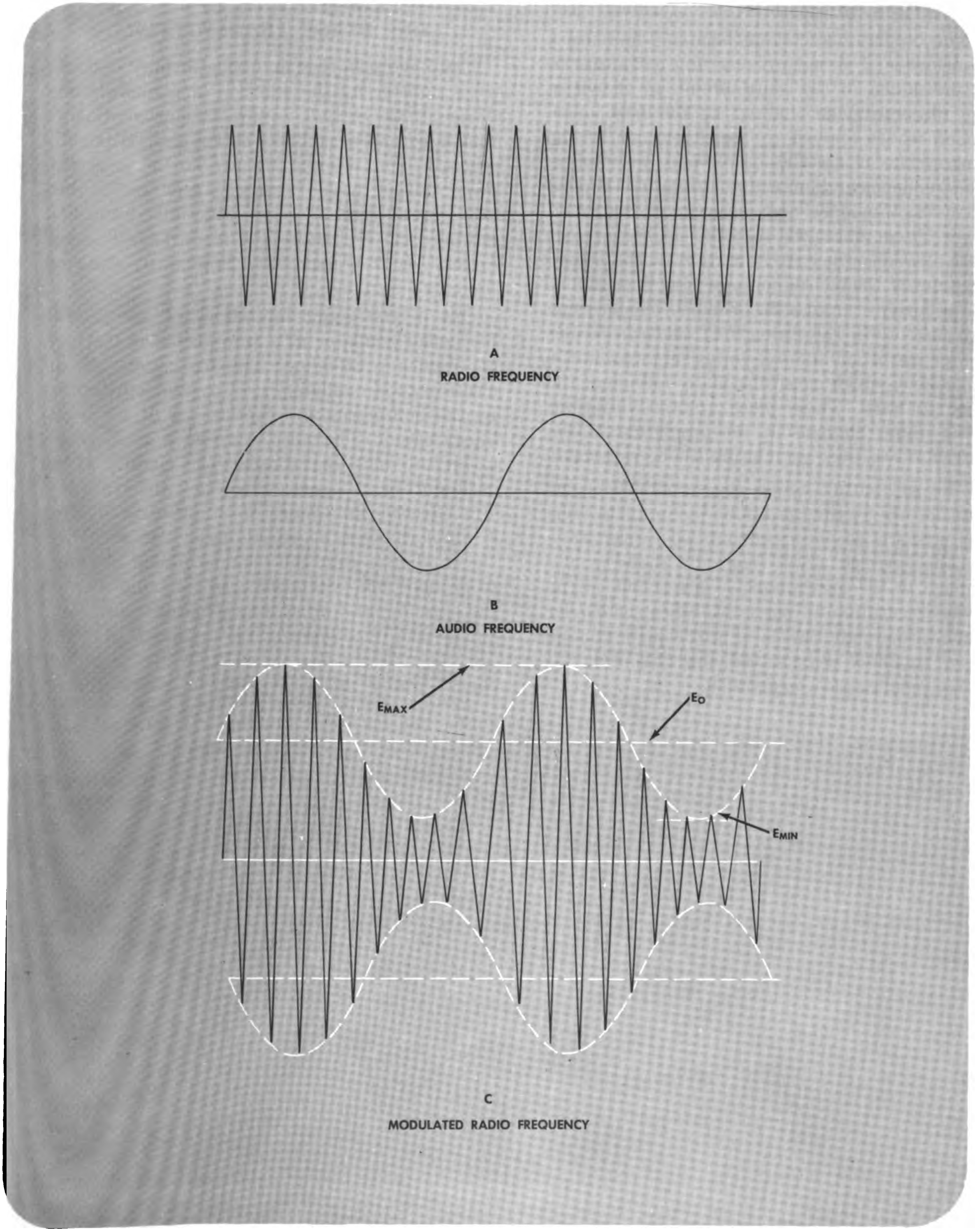
Interrupting the Continuous Wave



A Carrier Wave

A second method, known as *amplitude modulation* (AM), consists of increasing and decreasing the amplitude of the carrier wave in accordance with the changes in an audio (voice or music) signal. The waveforms on the next page show a carrier wave, an audio signal, and an amplitude-modulated carrier wave.

A third method, known as *frequency modulation* (FM), consists of varying the frequency of the carrier wave in accordance with the changes in an audio signal. The amplitude of the carrier wave is not varied as you can see in the illustration of the frequency-modulated carrier wave on page 254. *Phase modulation*, also called indirect frequency modulation, consists of varying the phase of the carrier wave in accordance with the changes in the audio signal without changing amplitude. In this chapter, only the principles of amplitude modulation are discussed.



Amplitude Modulated Carrier Wave

Modulating An Oscillator

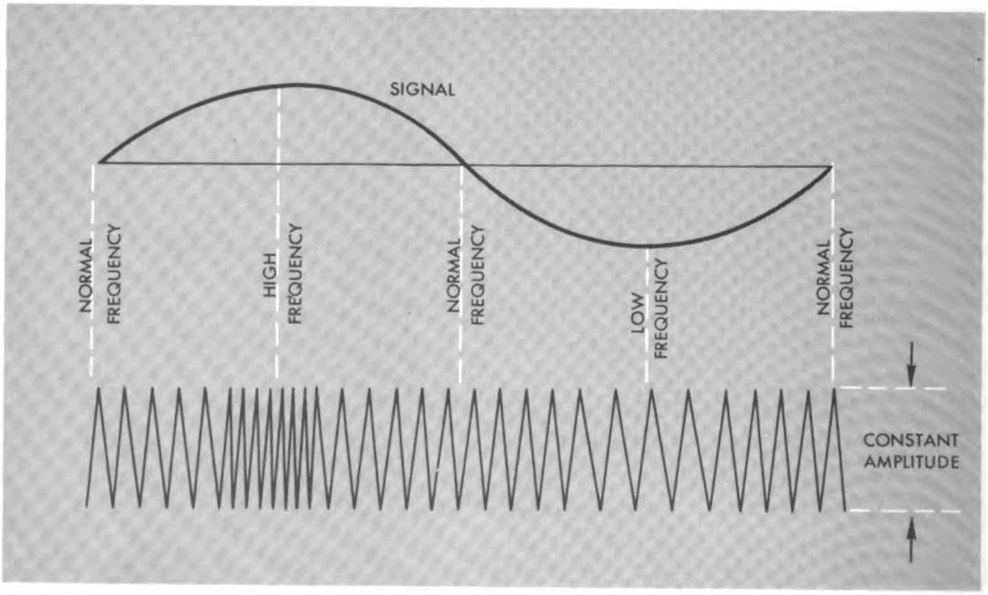
The oscillator circuit on page 250 becomes a transmitter if a microphone and a transformer are added in the plate circuit, as shown in the illustration at the top of the next page. Sound waves applied to the diaphragm of the microphone cause the diaphragm to vibrate. The frequency and intensity of vibrations depend upon the frequency and intensity of the sound wave. These vibrations vary the electrical resistance offered to the battery current by the carbon granules of the microphone, causing fluctuations in the current through the primary. The fluctuating currents in the primary of the transformer induce corresponding AC voltage changes in the secondary, which cause the plate voltage of the tube to vary. These plate voltage variations bring about changes in the amplitude of the carrier wave, thus modulating the carrier wave.

The AC output across the secondary of the transformer T_1 the variations in the plate voltage, and the amplitude-modulated carrier wave — all of which result from the sound waves applied to the microphone — are shown in the lower illustration on page 255. Since the carrier wave is modulated by causing the plate voltage of the oscillator to be varied, this type of modulation is called *plate modulation.*

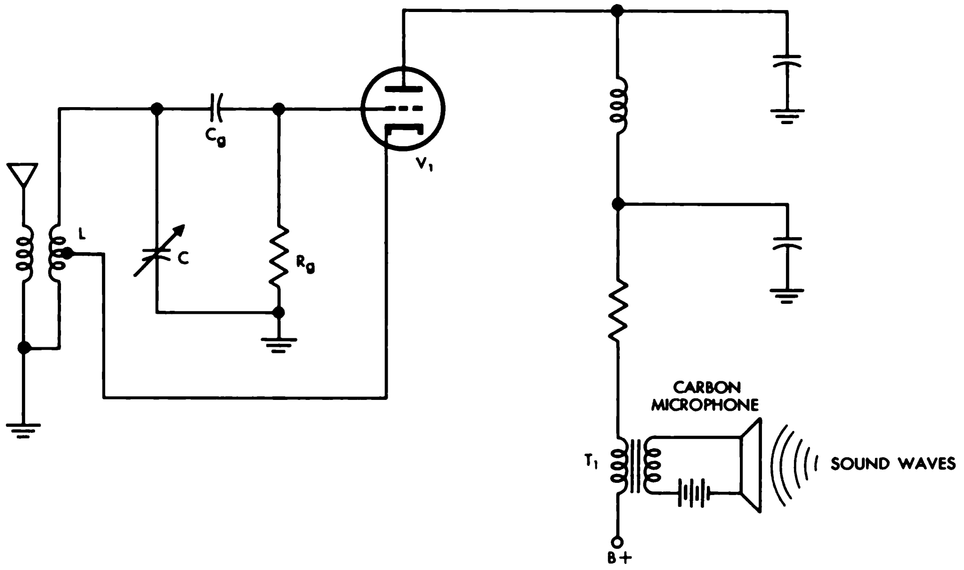
LIMITATIONS OF A ONE STAGE TRANSMITTER

You have seen that an oscillator may become a transmitter by the addition of a modulating device. In the early days of radio such oscillators were used as transmitters by coupling them directly to the antenna circuit. To increase the power output, larger tubes were used and higher voltages applied. For present day demands, however, an oscillator used alone is not satisfactory as a transmitter. Too many factors limit its usefulness.

One of the major limitations of a variable frequency oscillator as a transmitter is its lack of frequency stability. The frequency of an electron tube oscillator is, of course, controlled by the inductance and capacitance of its oscillatory circuit. Therefore, coupling a load (such as an antenna) to the oscillatory circuit of an oscillator has a great effect on its stability. The effective impedance of an antenna system is often changed by variations in the weather: the wind may shift the relation of the antenna to ground; rain, sleet, snow, or ice may fall on the antenna; and other weather conditions may change the impedance of the antenna. A varying load reflected into the oscillatory circuit will cause a change in the frequency of the oscillator.



Frequency-Modulated Carrier Wave



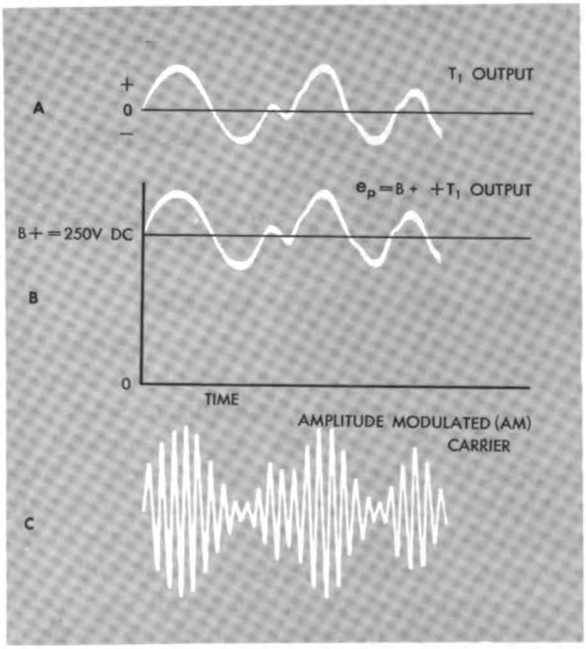
Making an Oscillator a Transmitter

In VHF and UHF circuits another kind of problem arises. Though the frequency stability required by such circuits can be attained by a crystal oscillator, the frequency range of a crystal is limited. In VHF and UHF, therefore, the crystal oscillator cannot stand alone. To provide the needed frequency, the crystal oscillator must be followed by frequency multipliers as intermediate stages in many transmitters. The usual VHF and UHF transmitter therefore contains an oscillator, one or more frequency multiplying stages, and a stage of RF amplification. Additional circuits are provided for keying or modulating the transmitter.

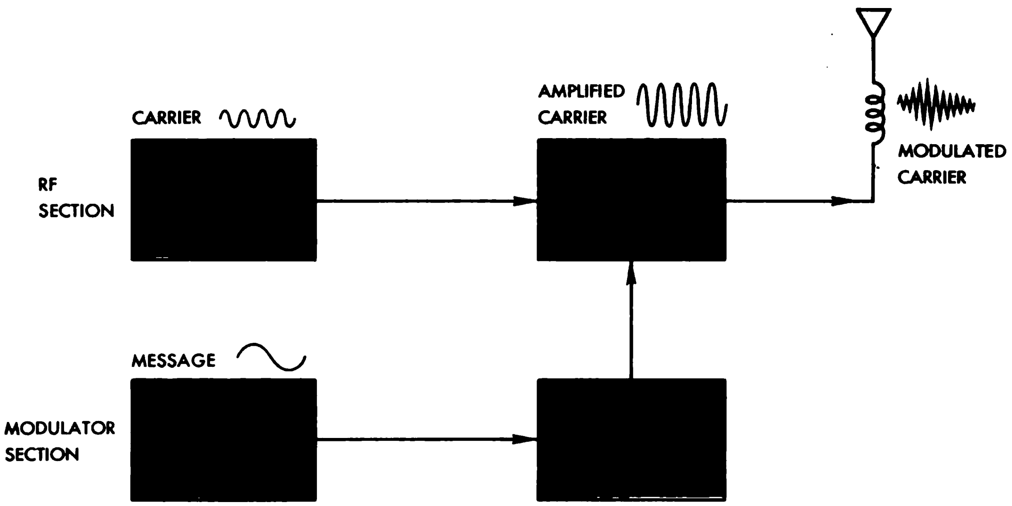
Another limitation of an oscillator as a transmitter is power output. To provide sufficient power for modern radio transmission, the oscillator must be followed by buffer and power amplifier stages which are discussed later in this chapter. With these stages, present day transmitters can use small tubes as oscillators even in high power transmitters. Tubes used as oscillators are often of the same type as those used in ordinary receiver output stages.

Operating efficiency is another problem of an oscillator as a transmitter. Coupling a load to the inductance of an oscillator circuit causes a decrease in the impedance of the circuit. This varying impedance in the tank circuit affects the voltage feedback ratio, as well as the frequency of oscillation. Changes in load value or coupling consequently reduce the efficiency of the oscillator.

Due to these disadvantages, the use of an oscillator by itself as a transmitter has be-



Effect of Sound Waves on Transmitter Output



Block Diagram of a Transmitter

come obsolete. The practical transmitter circuit today consists of a radio-frequency section and a modulation section. As you can see in the block diagram, the radio-frequency section consists of an oscillator and amplifier whose function it is to produce a carrier wave with sufficient power to be radiated over a given area. The modulator section consists of a microphone and audio amplifiers whose function it is to build up the audio signal so that it can be used to modulate the output of the RF section.

RADIO FREQUENCY SECTION

The simplest type of RF section for a transmitter in use today is a master oscillator-power amplifier (commonly called MOPA).

Master Oscillator Power Amplifier

The MOPA was developed to overcome the frequency instability of an oscillator coupled directly to an antenna. Its oscillator stage, isolated from the antenna, is not affected by changes in antenna-to-ground capacity. With a MOPA, if there are such changes, the only noticeable effect on the circuit is a change in the resonance conditions between the antenna and power amplifier stage. If the circuit is properly neutralized, the frequency stability of the oscillator will not be affected. The loss in radiation efficiency will be practically negligible.

Neutralization is a process of counteracting feedback which tends to make an RF amplifier operate as an oscillator. Since RF amplifiers are tuned, they tend to oscillate if feedback in proper phase occurs. It is therefore necessary to neutralize RF amplifiers if they are to perform their functions as amplifiers without generating oscillations. Neutralization is accomplished by feeding back to the input circuit a voltage which is 180° out of phase with any positive feedback which occurs through the interelectrode capacitances. For this purpose, a capacitor is usually used, and the voltage fed back exactly cancels any positive feedback that occurs.

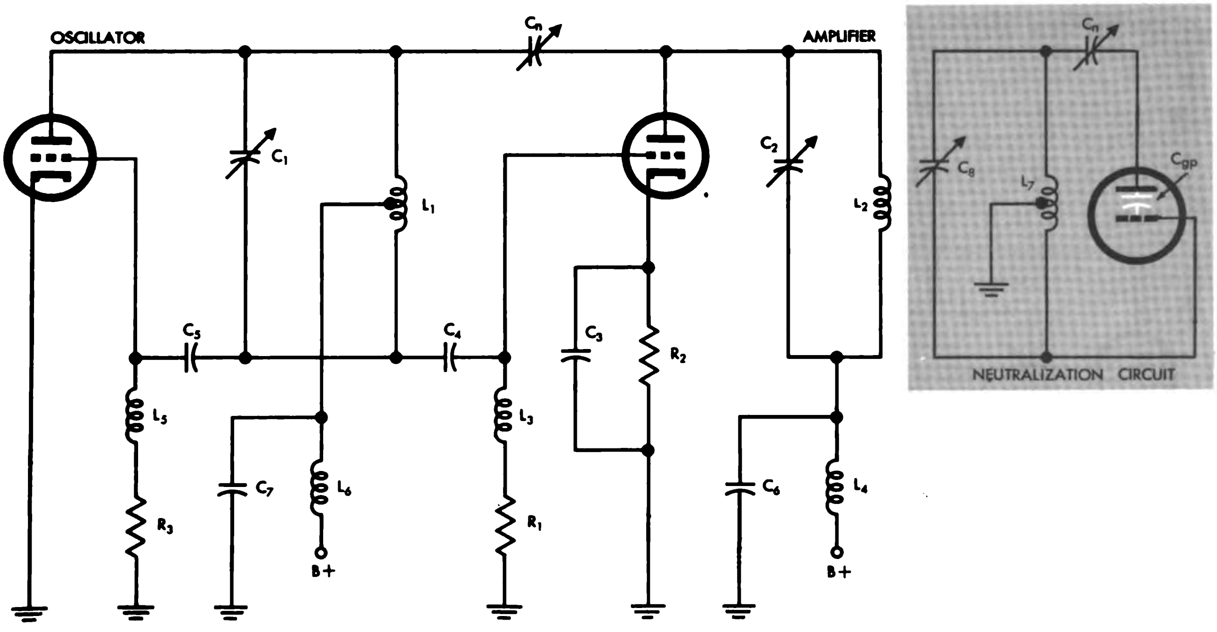
The reason that the oscillator of a MOPA transmitter is effectively isolated from changes in the antenna-to-ground capacitance is that the power amplifier circuit is located between the oscillator and the antenna. With the power amplifier neutralized, changes in its plate impedance will not be reflected into the grid circuit and back to the oscillator. The AC voltage applied to the grid of the PA from the oscillator will cause the plate tank of the PA to oscillate at the frequency of the oscillator. It is true that the antenna-to-ground capacitance is reflected back and becomes a factor in resonating the plate tank of the power amplifier. Still, the power amplifier is loaded and the Q of the circuit is low. This means that the plate tank has a broad resonant curve. Therefore, small changes in capacitance due to a changing load (a swinging antenna) will not much affect the tuning of the circuit. As a result, the grid signal will produce oscillations at its own frequency in the plate tank even when the values of L and C are slightly changed. Therefore, as long as the oscillator frequency is constant, the frequency of oscillations in the plate tank remain relatively unchanged.

OSCILLATOR CONSIDERATIONS. The MOPA may use any of the common variable frequency oscillator circuits, but usually not a crystal oscillator. If a crystal oscillator is used in a high frequency transmitter, an MOPA setup is not used. Instead, the crystal-controlled oscillator is usually followed by stages called buffers and frequency multipliers.

Selection of an oscillator depends on the purpose for which the transmitter is intended. For example, at frequencies up to approximately 500 kc, the Colpitts oscillator is generally chosen, for it is stable and dependable. At frequencies above 500 kc (in the medium frequency range) the Hartley circuit is chosen for its stability in that range.

Now examine the schematic diagram of the practical MOPA circuit shown below. It is complete with RF chokes and bypass capacitors. The oscillator is a series-fed Hartley circuit. Its output, coupled to the amplifier through C4, excites the amplifier grid circuit.

AMPLIFIER CONSIDERATIONS. The power amplifier also varies in design according to the purpose of the transmitter, but almost always operates class C. The output delivered to the antenna may vary from a few watts up to 50,000 watts, depending on the purpose of the transmitter. Purpose likewise, governs circuit considerations. The tube data charts outline typical operating conditions for RF amplifiers, specifying voltages, power output, required excitation voltages, bias, rated grid currents, and plate dissipation. Of these factors, power output is most important. A power-amplifier tube is selected and its stage designed for the desired output. The power of the power-amplifier is also a determining factor in the selection and design of the oscillator, since the oscillator must provide the proper amount of excitation to give the power amplifier its required output.



Master Oscillator Power Amplifier

The bias of the power amplifier must conform to the operating bias rated for the tube. When bias depends on grid current, safety bias is also used to prevent the tube from being damaged by excessive plate current if grid excitation should be removed. In practice, the amount of safety bias is somewhere between the value required to cut off the plate current and the value required for normal operation. In the schematic diagram of the MOPA, grid leak bias is developed across R1 and safety bias across R2.

Though the MOPA does not have to be neutralized each time it is tuned, it should be checked at specified intervals, because aging, handling, and temperature changes of circuit components over a period of time may make readjustment necessary. Neutralization in the MOPA shown above is accomplished when C_n is properly adjusted, since the charging current for C_n and C_{gp} is the same. Therefore no current flows through L7 to start parasitic oscillations.

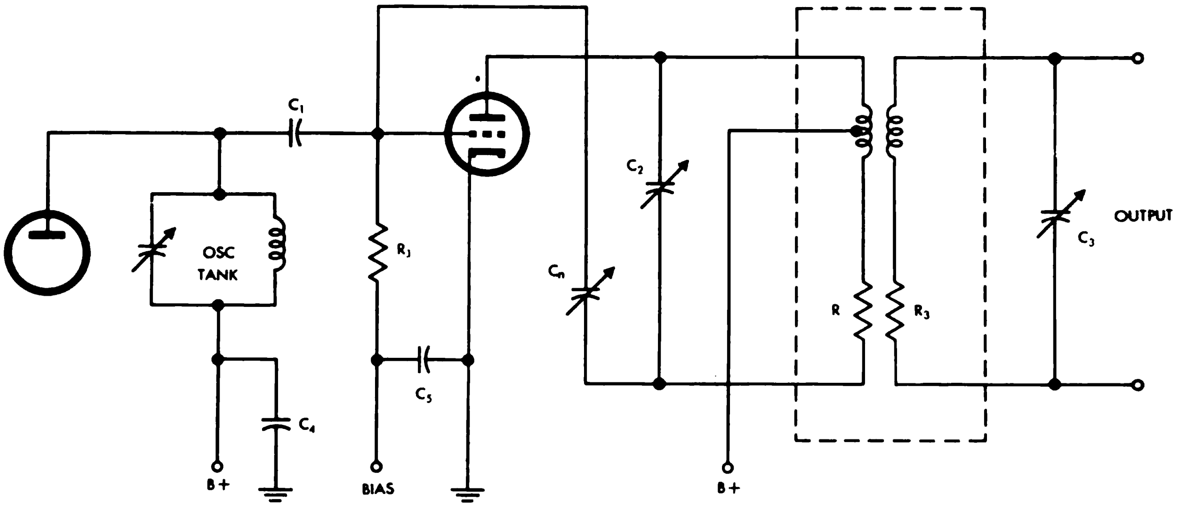
Buffer Amplifiers

To prevent interference to other radio services, radiation at harmonic frequencies must be kept at a minimum. To keep harmonic radiation down, many transmitters use crystal oscillators, which have relatively low frequency and low power. With the oscillations occurring at low power there is relatively little danger of interference. However, to raise the power output and frequency to the desired level, these crystal oscillators drive buffer amplifiers and or frequency multipliers. The various amplifier stages have different names. An amplifier with the same output frequency as the input frequency is known as a straight amplifier. When the function of such an amplifier is to isolate the oscillator from the other stages of the transmitter, it is called a buffer amplifier.

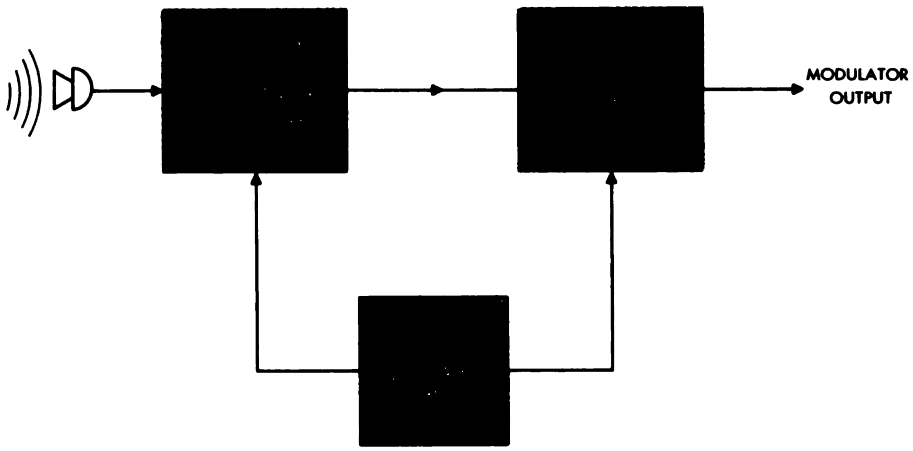
In the case of a MOPA transmitter where the PA is coupled directly to the oscillator, the oscillator is subject to a changing load when the PA is keyed. The input to the amplifier represents the load on the oscillator. Keying the power amplifier therefore has the effect of alternately connecting and disconnecting the load on the oscillator. This changing load causes the oscillator to become unstable. So a buffer amplifier is used. The purpose of the buffer is to act as an amplifier while isolating the oscillator from the power amplifier.

In the buffer amplifier circuit shown below, you can see that the buffer operates essentially like any other voltage amplifier. The RF voltage developed across the oscillator tank is applied between grid and cathode of the buffer tube through coupling capacitor C_1. C_5 is included to bypass the RF voltage around the bias supply. Resistor R furnishes grid leak bias for the buffer.

The buffer plate load may be a tuned circuit as shown, or it may be the primary of an air-core transformer. The tuned circuit offers a



Buffer Amplifier



Block Diagram of Modulator Section

fairly constant load impedance and good response over a band of frequencies. Once the circuit is adjusted with the values of capacitors C_2 and C_3 fixed, the load remains unchanged.

The RF section of a transmitter (see block diagram on page 256) usually has one or two stages of voltage amplification between the oscillator and the RF power amplifier.

MODULATOR SECTION

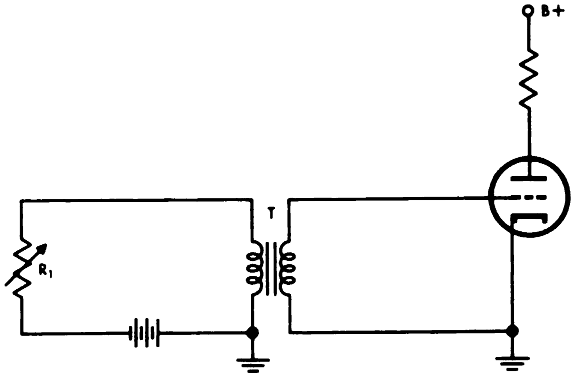
A microphone, voltage amplifiers, and power amplifier make up the modulator section of a transmitter. The audio amplifier stage before the power amplifier is sometimes called the *driver*, and the audio power amplifier whose output modulates the carrier is sometimes called the *modulator*. Study the block diagram of the modulator section to see the relationship between the stages.

Microphone

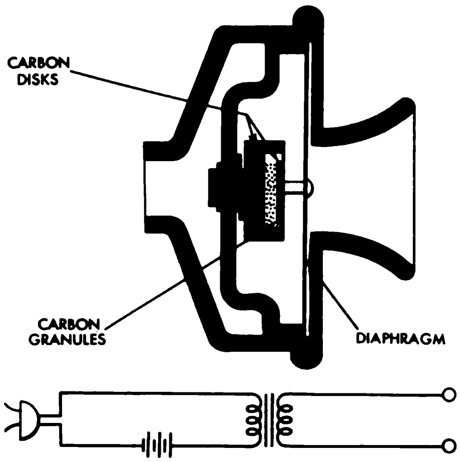
The illustration to the right shows the primary circuit of a transformer consisting of the primary winding, a variable resistor and a battery in series. The secondary winding is coupled into a triode. Notice that as the variable resistor is varied the current flowing in the primary is changed, causing a voltage to be induced into the secondary.

The carbon microphone operates in essentially the same manner. It consists of carbon granules loosely packed between two copper disks as shown in the top illustration on page 260. As the sound waves strike the diaphragm they cause the carbon granules to compress (resistance decrease) or to expand (resistance increase) in accordance with the pressure exerted by the sound wave.

If the microphone replaces R_1, of the schematic below, the current in the primary of the transformer will vary at the rate of the impressed sound wave. The voltage induced in the secondary is then coupled into the grid of the voltage amplifier tube.



Generating a Varying Voltage by Varying Resistance



Carbon Microphone and Schematic

Voltage Amplifier

The schematic diagram of a typical modulator section is shown below. The input to the amplifier is developed by the carbon microphone. The low voltage supplied to the microphone must not be too high. Otherwise, the average microphone current will tend to make the carbon granules pack together. Excessive microphone current also causes noise.

The voltage supply is bypassed by C_1 so that no audio frequency voltage is dissipated by the voltage source. The microphone transformer T_1 has a primary impedance of about 100 ohms and a secondary impedance of 120,000 ohms to give a relatively large voltage gain. The secondary of the microphone transformer is shunted by the input level control potentiometer R_1. This control is normally adjusted for 100 percent modulation on signal peaks as observed on the screen of an oscilloscope. Percentage of modulation will be discussed later in this chapter. Cathode bias for this stage is supplied by R_2 and C_2. The signal from the plate of the voltage amplifier is RC coupled to the following stage.

The output of the amplifier stage may be applied to a driver stage or directly to the power amplifier as shown.

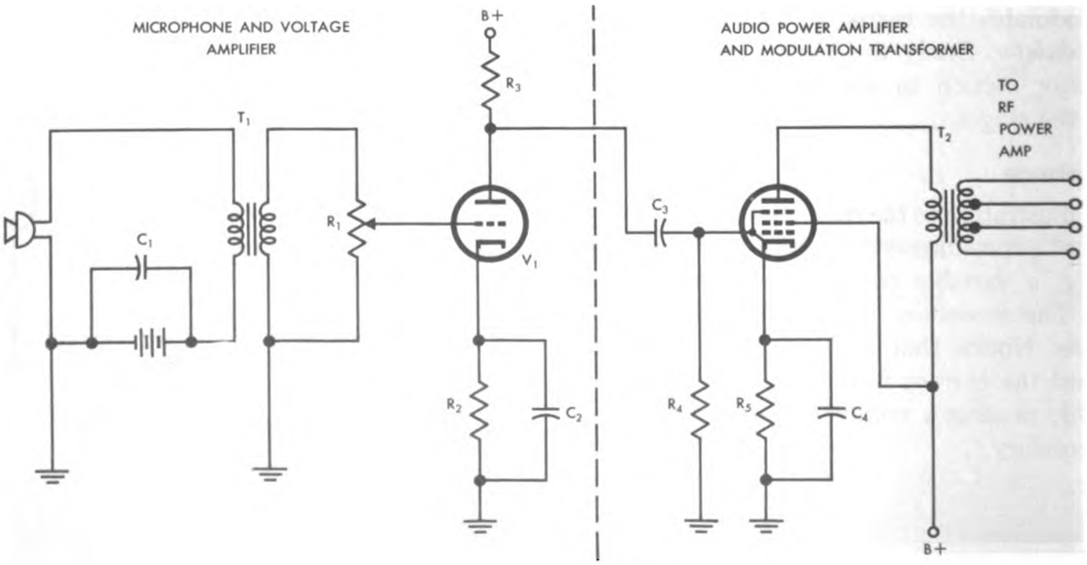
Audio Power Amplifier

The final stage of the modulator section is the audio power amplifier. In the schematic shown, this stage is an RC-coupled, class A amplifier.

Cathode bias is obtained through R_5 and C_4. B+ is applied through the primary of the output transformer.

The output transformer is known as the modulation transformer. It must match the impedance of the audio power amplifier to the impedance of the RF power amplifier. For this reason, the secondary is usually tapped at different points. In addition, the modulation transformer must have a high current rating. Its primary winding must be able to carry the plate current of the audio power amplifier, and its secondary winding must be able to carry the plate current of the RF power amplifier.

Typical Modulator Section


MICROPHONE AND VOLTAGE AMPLIFIER
B+
R3
T1
R1
C1
V1
R2
C2
C3
AUDIO POWER AMPLIFIER AND MODULATION TRANSFORMER
TO RF POWER AMP
T2
R4
R5
C4
B+

PRINCIPLES OF AMPLITUDE MODULATION

An amplitude-modulated wave and the two waves which combine to produce it are shown in the illustration on page 262. The waveform at A represents the unmodulated RF voltage output of a transmitter. Its frequency is determined by the oscillator portion of the RF section, and its amplitude and power are determined by the amplifier stages of the RF section. The waveform at B represents an audio frequency sine wave used to modulate the carrier.

Waveform C is different from both wave A and wave B. It represents the resultant when the two waves are combined. As the audio-frequency wave swings positive, it increases the amplitude of the carrier wave from E_o to E_{max}; when it swings negative, it decreases the amplitude of the carrier wave from E_o to E_{min}. Therefore, the increase and decrease of the carrier wave occur at the frequency of the modulating signal. The modulation shape (area between dotted sine waves) which results is called the *modulation envelope*. This modulation envelope represents the intelligence transmitted by the carrier.

The waveform at C shows two frequencies: the carrier frequency and the modulating frequency. However, there are at least two other frequencies present in the amplitude-modulated waveform. How this happens may be made clearer by an analogy using sound waves. When two musical notes of slightly different pitch are sounded at the same time, a third note is heard. This new note has a frequency that is equal to the difference between the other two tones and is called the difference tone, or *beat frequency*.

A similar effect occurs in an electrical circuit. When two voltages of different frequencies are combined, a voltage is produced whose frequency is equal to the difference of the frequencies of the applied waves; this new frequency is also referred to as the *difference frequency*. A voltage of another frequency is also produced — one whose frequency is equal to the sum of the frequencies of the applied voltages. The frequency of this wave is sometimes called the *sum* frequency. The process of producing sum and difference frequencies by combining two or more frequencies is known as *heterodyning*.

The sum and difference frequencies that are the result of heterodyning are called *sidebands* or *sideband frequencies*. For example, if the carrier frequency is 1000 kc and the modulating frequency is 1000 cps, the upper sideband is 1001 kc and the lower sideband is 999 kc. The presence of these sideband frequencies can be demonstrated. If a too highly selective receiver is tuned to the carrier frequency, the sidebands are not received. Consequently, no modulation is heard. This is why tuned circuits in a receiver must have a bandpass sufficiently wide to amplify the sidebands as well as the carrier frequency, as explained in the study of amplifiers. The bandpass of these tuned circuits in a communications receiver is approximately 6000 cps wide (3000 cps above and below the carrier). In a broadcast receiver, the bandwidth should pass all the sideband frequencies, from the highest to the lowest, equally well. Otherwise, fidelity will be reduced.

Percentage of Modulation

The percentage of modulation (M) depends on the relationship between the amplitude of the modulated wave and the amplitude of the unmodulated wave. It is most conveniently expressed by the formula,

$$M = \frac{E_{max} - E_{min}}{2E_o} \times 100$$

Where E_{max} is the maximum amplitude of the modulated wave, E_{min} is the minimum amplitude of the modulated wave, and E_o is the amplitude of the unmodulated wave. (See waveform C on page 262.

If the maximum amplitude of the modulated wave is 10 volts, the minimum amplitude of the modulated wave is 0 volts, and the amplitude of the unmodulated wave is 5 volts, then

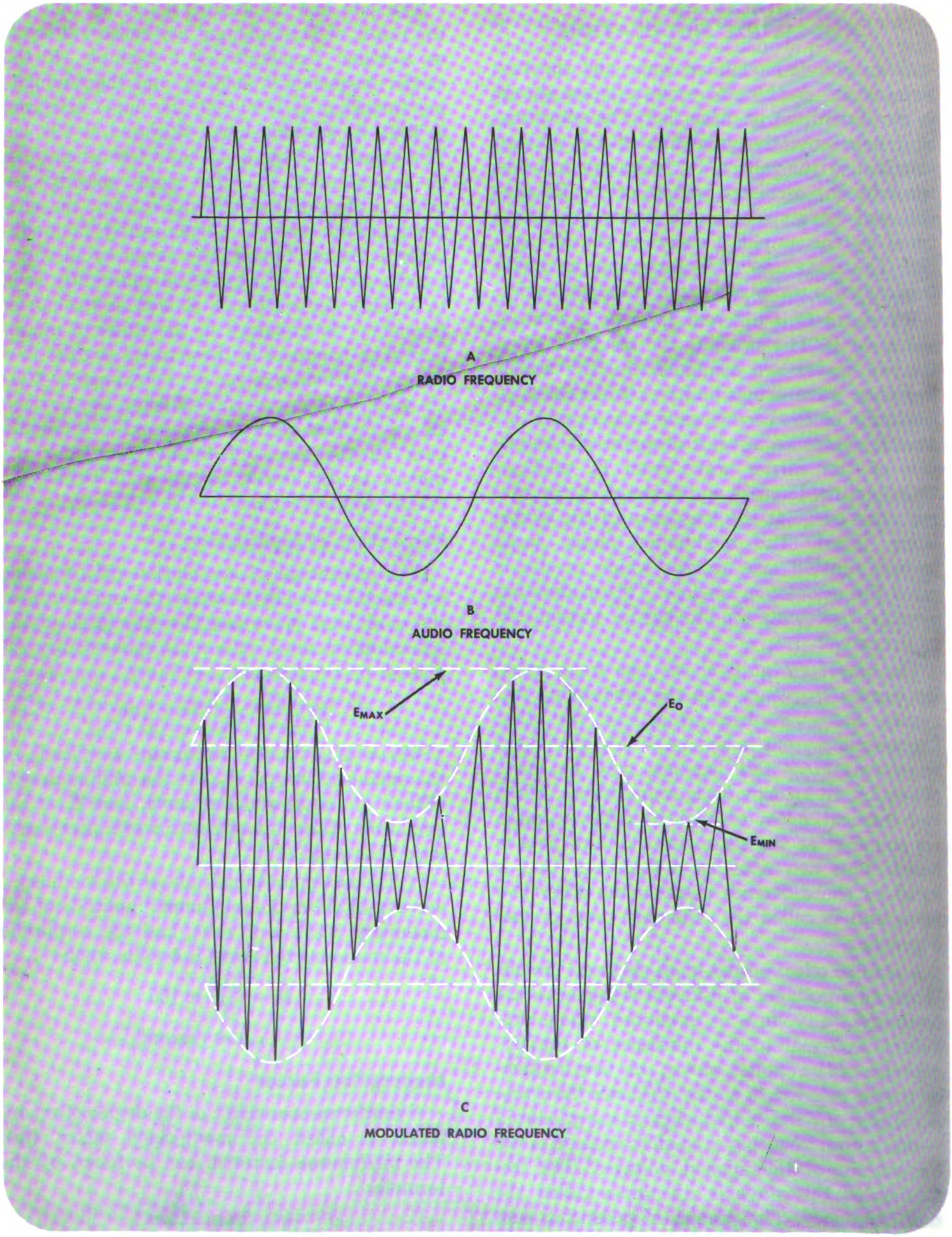
$$M = \frac{10-0}{10} \times 100$$
$$= 100 \text{ percent}$$

Modulation Factor

A similar measurement of the degree of modulation is called the modulation factor (m). The modulation factor m can be found by the formula:

$$m = \frac{E_{max} - E_{min}}{2\,E_o}$$

Substituting the above values, the modulation



Amplitude Modulation Waveforms

factor is $\frac{10-0}{10}$, or 1. A modulation factor of 1 corresponds with a modulation percentage of 100 percent.

The greater the percentage of modulation, the greater is the amplitude of the sidebands and the greater is the distance coverage of the transmission. If the percentage of modulation is low, little power appears in the sidebands and the distance coverage of the transmission is likewise small. On the other hand, if the percentage of modulation is high, there is more power in the sidebands and the distance coverage is greater. However, the percentage of modulation should not exceed 100 percent. If it does, severe distortion of the intelligence carried by the wave will result.

The formulas just presented are only generally useful. Each should really be two formulas, since it is very difficult to design a circuit that will always produce the same degree of modulation on both the positive and the negative peaks of the audio modulating voltage. The following formula gives the degree of modulation on the positive peaks. (The terms are the same as in waveform C on page 262.

$$\frac{E_{max}-E_o}{E_o}$$

The expression for finding the degree of modulation on the negative peaks is.

$$\frac{E_o-E_{min}}{E_o}$$

Since the minimum amplitude of the modulated wave cannot be less than zero, the degree of modulation cannot exceed one, or unity, on the negative peaks. However, the degree of modulation can exceed unity on the positive peaks. This condition is called over-modulation. It causes abnormal distortion of the carrier wave, creating sideband frequencies far enough above and below the carrier frequency to interfere with received signals from transmitters operating on adjacent commuincations channels.

METHODS OF AMPLITUDE MODULATION

There are many methods of obtaining an amplitude-modulated wave. No attempt will be made here to discuss them all; however, some of the more common methods will be explained. All methods of modulation involve nonlinear impedances in which the graph of the current against the applied voltage is not a straight line. Most methods involve the use of suitable vacuum tubes and associated circuit components. These vacuum tubes provide the necessary nonlinear impedance. Other factors to be considered are the amount of power to be handled, the percentage of modulation desired, and the cost of the necessary equipment,

The most common method of amplitude modulation is *plate modulation* of the RF power amplifier. This type of modulation is called *high level* modulation because it requires a large amount of audio power.

Modulation voltage can also be applied to the control grid, cathode, screen grid, or suppressor grid of the RF power amplifier or of a previous RF stage. This type of modulation is called *low level* modulation because it requires little audio power.

Only plate and grid modulation are described here, since you will encounter these methods of modulation most frequently.

Plate Modulation

A high-frequency oscillator may be used to produce a modulated wave by varying the voltage applied to its plate at the desired modulating signal rate, as described on page 254.This method of producing an amplitude-modulated wave is almost obsolete. The varying plate voltage causes variations in frequency which are not desirable. Since the tube may cease oscillating at low values of plate voltage, this method makes it difficult to obtain a high degree of modulation.

The disadvantages can be avoided if the modulation is accomplished in the plate circuit of a class C amplifier instead of in the plate circuit of an oscillator. A simplified circuit of this type is shown in the illustration on page 264.V_1 is an audio amplifier which amplifies the modulating signal. As the plate voltage of V_1 is varied by the signal, the voltage applied to the plate of V_2 varies at the same rate. The RF oscillator generates the carrier wave which is fed to the grid of the class C amplifier.

With no signal applied, the plate current of the class C amplifier appears as shown in wave A. The voltage in the output remains constant in

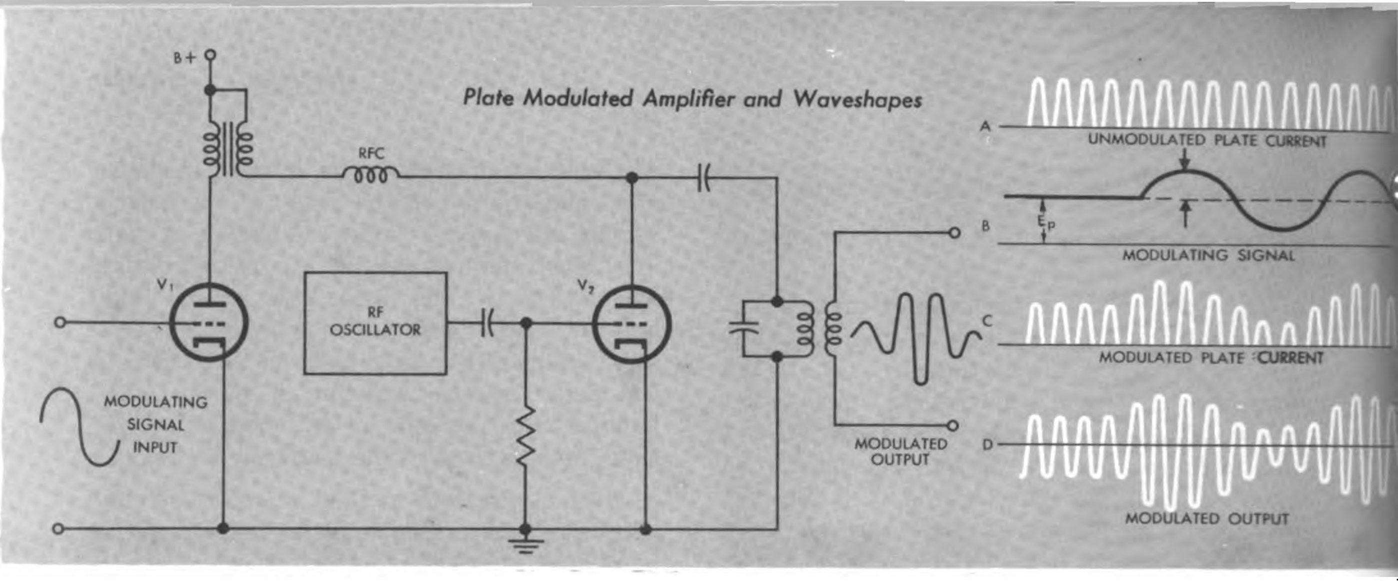


Plate Modulated Amplifier and Waveshapes

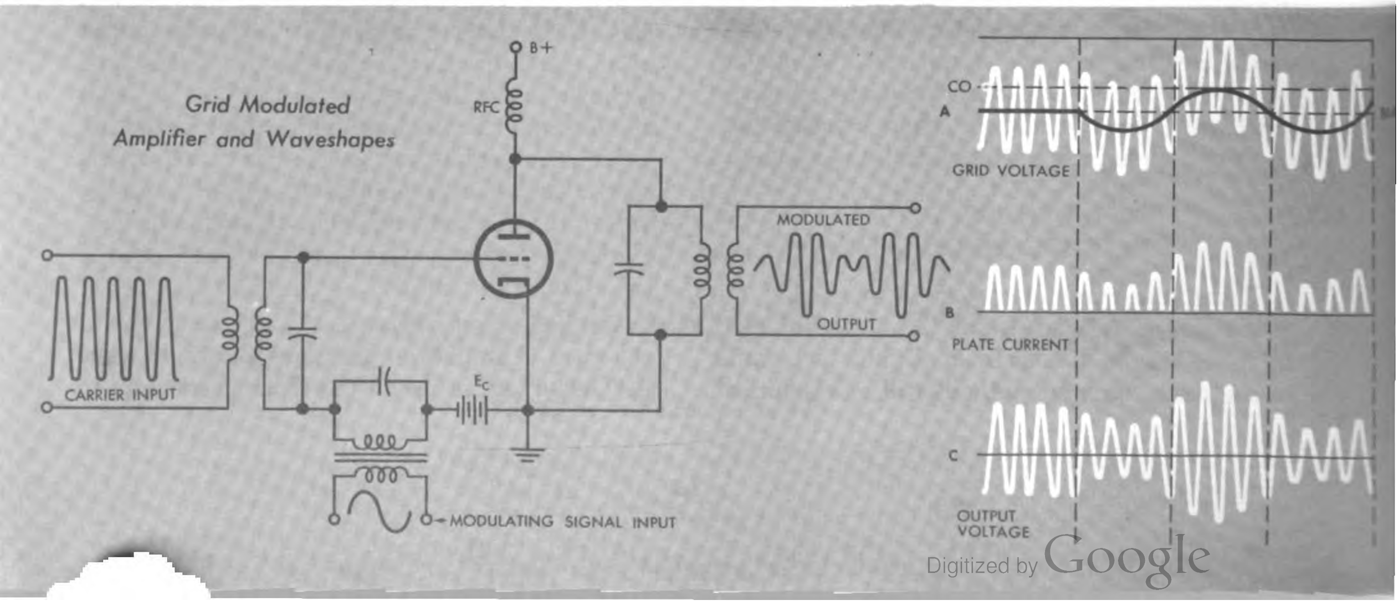
amplitude in the absence of a modulating signal, as shown in the first portions of waves C and D. When the modulating signal is applied, as shown in wave B, it causes a variation in the applied voltage. This variation in applied voltage causes the plate current to vary as shown in wave C. Because of the flywheel effect, the variations of plate current supplied to the resonant LC circuit cause AC variations in the voltage drop. This variation is illustrated by wave D, the modulated output. It contains the carrier and the sideband frequencies. These sideband frequencies are the result of impressing two different signals on a nonlinear impedance — the class C amplifier.

Grid Modulation

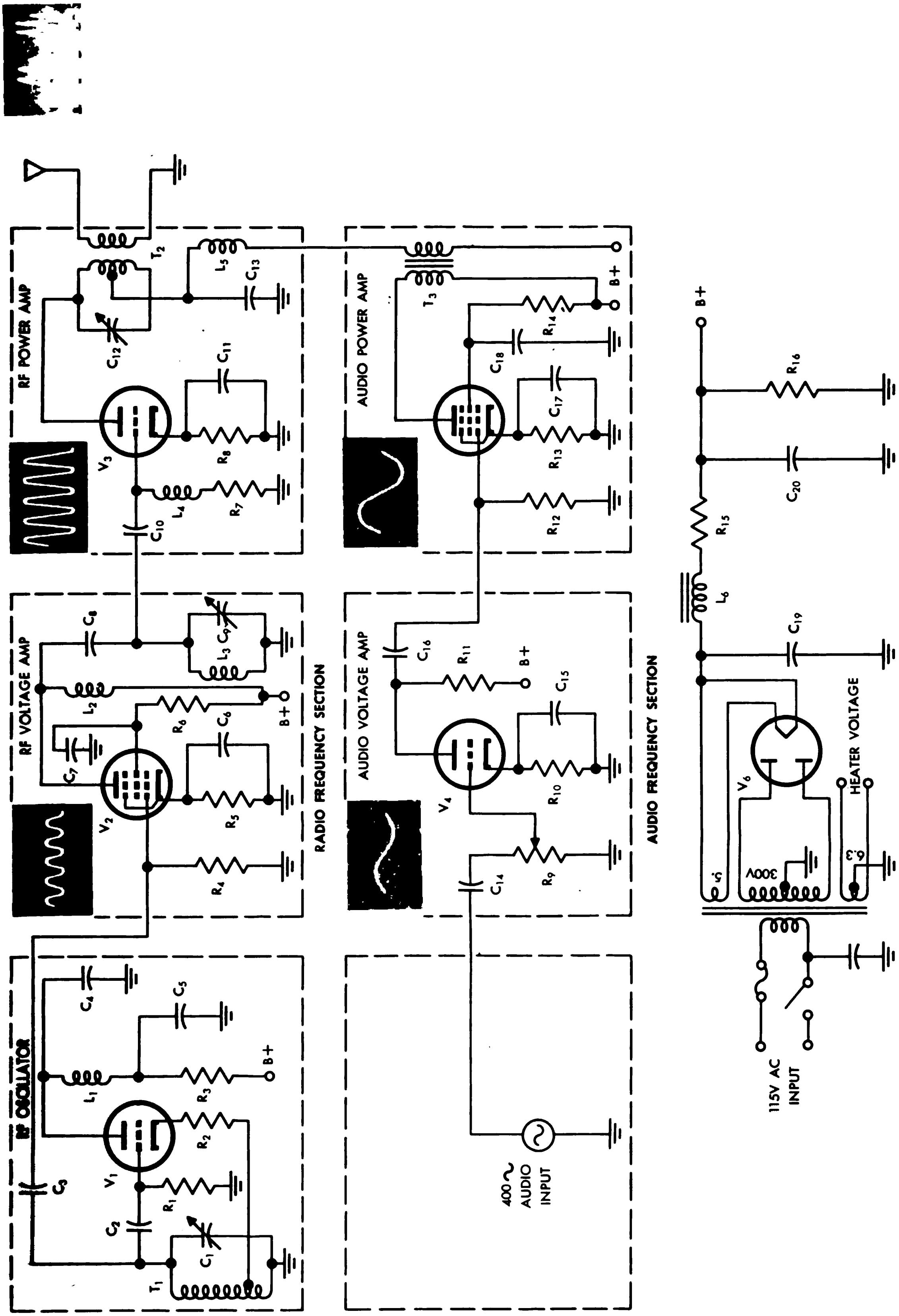
A grid-modulated amplifier is one in which the modulating signal is impressed upon the carrier in the grid circuit. A simplified circuit of a grid-modulated amplifier, together with its waveshapes, is shown below. The grid voltage of the amplifier is caused to vary at the rate of the signal frequency. This variation in grid voltage causes a variation in the power output of the amplifier, producing a modulated signal. Waveshape A shows the voltage applied to the grid of the amplifier. The amplifier is biased class C by means of a DC bias supply. The signal voltage causes the grid voltage to vary at the rate of the signal. The amplitude of the RF carrier varies with the modulating signal. Plate current flows only when the voltage on the grid is above cutoff, as shown in wave B. The average plate current varies in accordance with the modulating signal. This variation in plate current is fed to the resonant tank circuit, causing it to produce a modulated wave as shown at C.

Grid modulation has the advantage of requiring much less modulator power for 100% modulation than plate modulation. However, when grid modulation is used, it is difficult to prevent distortion of the modulation envelope.

Distortion can be kept to a minimum by careful regulation of the bias supply of the modulated RF amplifier and of the plate supply of the preceeding RF amplifier.



Grid Modulated Amplifier and Waveshapes



A Complete Radio Transmitter

A COMPLETE RADIO TRANSMITTER

You have seen (page 256) that the practical transmitter of today consists of a radio-frequency section and a modulator section. Each of these sections has been discussed separately. Study the illustration on page 265 where both of these sections have been combined and a power supply added to form a radio transmitter.

The power supply shown here is a full-wave rectifier with a Pi-type filter which supplies B+ for all tubes. In some instances more than one power supply may be used. Heater voltages are obtained from secondary windings of the power transformer.

The radio-frequency section is composed of a Hartley oscillator, an RF voltage amplifier, and an RF power amplifier which is plate-modulated. Note that the output of the oscillator is RC coupled to the voltage amplifier. The RF power amplifier is double tuned, with tuned circuits in the grid and the plate.

The modulator section shown uses an audio signal generator to generate a 400 ∼ input. This replaces the microphone and audio amplifier stage of the previous circuit. The audio output is amplified by an audio voltage amplifier and audio power amplifier. Both stages are RC coupled and operated class A. The output of the audio power amplifier is coupled into the plate circuit of the RF power amplifier through transformer T_3. The audio variations cause variations in the plate voltage of the RF power amplifier, thus modulating its output. The waveshapes for each of the stages are shown in the illustration.

Note how the amplitude-modulated carrier is coupled to the antenna circuit through transformer T_2 and then radiated into space. This radiated wave carrying the information with which it is modulated is picked up by the antenna of a receiver. The function of the receiver is to separate the information from the carrier so that it can be reproduced. How the receiver accomplishes this function is explained in the next chapter.

Chapter 29

Principles of Radio Reception

The receiver has the job of completing the communications cycle started by the transmitter. It must intercept some of the RF energy radiated by the transmitter; it must separate the intelligence from the RF carrier; and it must reproduce the intelligence as energy in some meaningful form such as sound, light, or mechanical energy. What a receiver is, and how it performs its function, is discussed in this chapter.

GENERAL CHARACTERISTICS OF RECEIVERS

Receivers may be described in terms of certain general characteristics such as *frequency coverage*, *sensitivity*, *selectivity*, and *fidelity*.

Frequency Coverage

The *frequency coverage* of a receiver is the range of RF which it can handle. Frequencies may be radiated anywhere from 10 kc to above 30,000 mc. However, no one practical receiver has yet been constructed that will handle this entire range of frequencies successfully. This range constitutes too big a chunk of frequencies. Even talking about it clearly requires that it be divided, as shown in the frequency spectrum chart at the top of the next page.

Another kind of division is shown in the frequency allocation chart also shown on the next page. This chart is based on the purpose for which the frequencies are used. A receiver might handle a complete band or only part of a band. At most, it will handle only two or three bands. Thus, the broadcast receiver has a frequency coverage of 550 kc to 1600 kc (a complete band), while some receivers used in Air Force communications have a frequency coverage of 225 mc to 399.9 mc, only part of a band. Most radar receivers operate in a small portion of the SHF (super high frequency) band.

Sensitivity

The *sensitivity* of a receiver is the measure of its ability to intercept weak signals and extract intelligence from them.

RADIO FREQUENCY SPECTRUM

Frequency Band		Designation	Abbreviation
KC	MC		
10-30		Very Low	VLF
30-300		Low	LF
300-3,000		Medium	MF
3,000-30,000	3-30	High	HF
30,000-300,000	30-300	Very High	VHF
300,000-3,000,000	300-3,000	Ultra High	UHF
3,000,000-30,000,000	3,000-30,000	Super High	SHF

A receiver intercepts many radiations. These radiations may originate from transmitters operating anywhere within the radio frequency spectrum. They may originate from transmitters located anywhere in the world. They may also originate from transmitters of widely varying power output. Thus, the RF radiations intercepted by a receiver are of widely varying signal strength. A highly sensitive receiver does a good job of completing communication, whether signals are weak or strong.

Selectivity

Selectivity is the measure of a receiver's ability to intercept a desired signal and extract its intelligence to the exclusion of others. From the point of view of a receiver, the RF radiation that comes along is composed of a hodgepodge of various frequencies and various signal strengths. A highly selective receiver must be able to concentrate on one signal and reject all others, even those whose frequencies are close to the frequency of the desired signal.

Fidelity

While sensitivity and selectivity are the measure of radio receivers' ability to intercept a weak signal and to extract the intelligence from that signal to the exclusion of all others, *fidelity* is the measure of receivers' ability to reproduce the intelligence of the signal. A receiver designed primarily for entertainment requires a high degree of fidelity. It must reproduce faithfully the sounds and sights on which entertainment depends — for example, the music of a symphony orchestra, or the fine details of a television picture. The fidelity requirements for simple communication are not so high. Still, reproduction must be faithful enough to make the message intelligible. This means reproduction without undue distortion. Voice reproduction must be understandable.

Physical Requirements of a Receiver

The components of a receiver are usually smaller and more compact than those of a transmitter. This is due chiefly to the fact that the power requirements of a receiver are less. As a result, power supply components are usually included right in the receiver unit.

FREQUENCY ALLOCATION

Frequency Band	Wavelength (in meters)	Uses
400 mc	0.75	Experimental
400 mc 106 mc	2.83	Government, Aircraft, Police, Television
108 mc 88 mc	3.41	Frequency Modulation
88 mc 50 mc	6.82	Television
50 mc 1600 kc	187.5	Ship-to-Shore, Aircraft, Police, Foreign, Government, Point-to-point, Experimental
1600 kc 550 kc	545.45	Commercial broadcast
550 kc 20 kc	15,000	Government, Commercial, Maritime, Ship-to-Shore, Aircraft, Point-to-Point, High Power Government, Transoceanic

BASIC REQUIREMENTS FOR RADIO RECEPTION

A simple practical receiver performs four essential operations — interception, selection, detection, and reproduction. By interception and selection, one RF signal is extracted from a multitude of RF signals. By detection, the intelligence is separated from the RF carrier. By reproduction, the intelligence is turned into some type of energy which reproduces the intelligence. Each of these operations, with the exception of selection, is shown in the illustration below.

Antenna

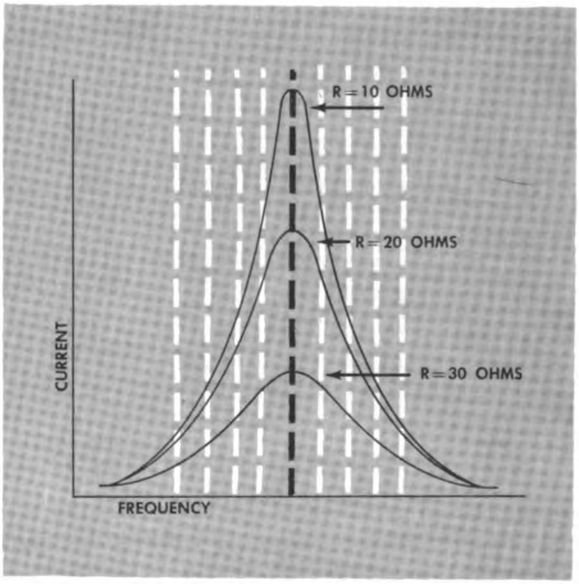
Actually an antenna in itself is a tuned circuit. It provides some selectivity. The antenna equivalent circuit in the illustration at lower right shows how an antenna can be a resonant circuit. L_a represents the distributed inductance of the antenna. C_a represents the distributed capacitance of the antenna. R_a represents the resistance of the antenna. The resonant frequency of an antenna can be determined by the formula,

$$F\ (\text{mc}) = \frac{159}{\sqrt{LC}}$$

where L is in microhenries and C is in micromicrofarads. Substituting typical values ($L_a = 50\mu h$ and $C_a = 200$ mmf) in the formula,

$$F\ (\text{mc}) = \frac{159}{\sqrt{(50)\ (200)}}$$
$$= 1.59 \text{ mc}$$

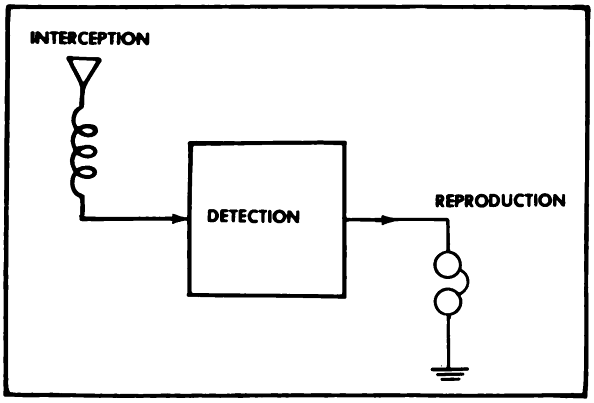
Note that frequency may be changed by varying L or C or both. When the antenna is a wire or a mast, L and C vary directly with antenna length. The longer the antenna, the lower the resonant frequency.



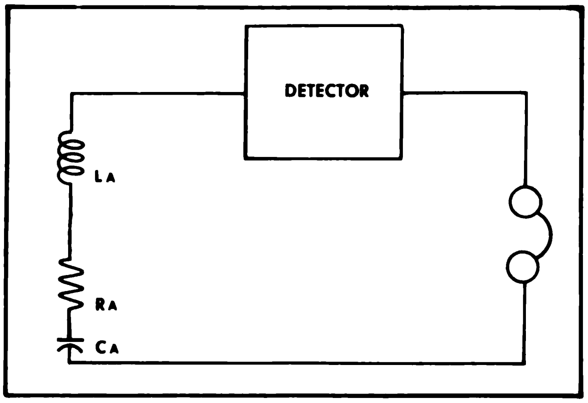
Effect of Resistance on Selectivity of an Antenna

The resistance of the antenna affects the sharpness of response of the tuned circuit. This is illustrated in the response curves above, which show the relative values of antenna current for varying values of resistance, with L and C remaining constant. Note that the smaller the resistance, the greater the response and the selectivity of the antenna.

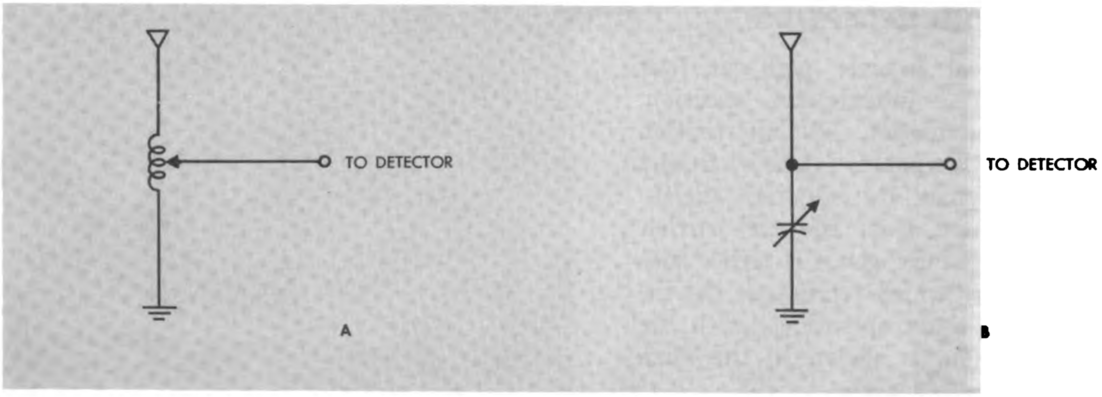
An antenna can be tuned in construction by making its length correspond to a quarter wavelength or a multiple of a quarter wavelength



The Simple Receiver



Antenna Equivalent Circuit

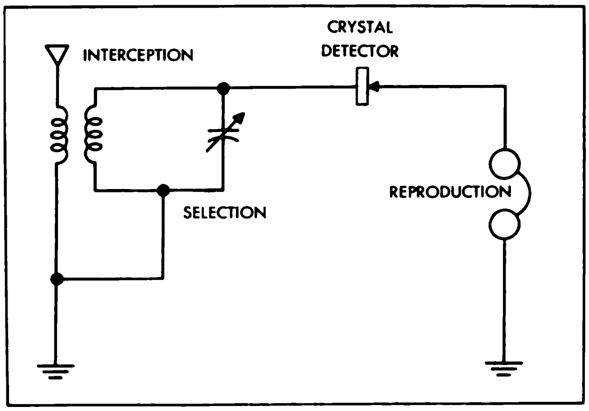


Varying Electrical Length of an Antenna

of a desired frequency. An antenna may also be provided with a means for varying its length for operational purposes.

At high frequencies, the selectivity provided by antenna tuning can be very critical and very important. At those frequencies, antennas are short, and their resistance is kept as small as possible. A slightly detuned antenna may make a receiver inoperative.

At low frequencies, antennas are long, and reduction of internal resistance is less practical. Here, antenna tuning is usually much less critical. The gain provided in a receiver by amplifier stages far surpasses the gain obtained by antenna tuning. Antenna length is of little importance in broadcast receivers. However, in simple receivers such as the one under consideration, there are no amplifier stages to provide signal amplification. Here, antenna length is important both for the sensitivity and the selectivity of the receiver.



Simple Tunable Receiver

The receiver antenna intercepts the signal by being in the path of the radiated RF. When magnetic lines of force cut a conductor (the antenna) an EMF is induced in the conductor. This induced voltage reproduces the RF carrier and any frequency or amplitude variations with which it is modulated. The induced voltage is greatest when the antenna is tuned to the frequency of the RF carrier. The induced voltage represents the intercepted energy needed for the first steps in the receiving process — interception and selection.

Tuning for Selectivity

As stated before, the resonant frequency of an antenna can be changed by varying antenna inductance, antenna capacitance, or both. This can be done by placing a variable inductance or a variable capacitance in the antenna circuit as shown in A and B of the illustration above. Adding an inductance or capacitance has no effect on the physical length of an antenna, but it has considerable effect on the electrical length. It is the electrical length which determines the resonant frequency of the antenna. When the added inductance or capacitance is variable, the electrical length of the antenna can be changed to achieve exact tuning for a desired signal.

Higher selectivity can be obtained by coupling the received signal into a separate tuned circuit, as shown in the illustration to the left. With such a selector circuit added, selection may be regarded as a separate step, distinct from antenna tuning. The antenna may then be regarded as being primarily for interception.

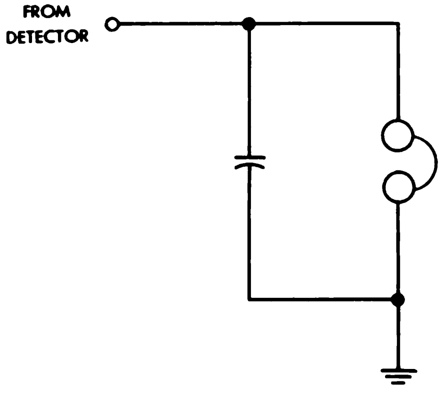
The use of a transformer improves both sensitivity and selectivity. The transformer steps up the voltage and isolates the tuned circuit from the resistance of the antenna. For tuning, either the capacitor or the coil could be made variable.

Detection

The process of separating the intelligence from the modulated RF signal is called *detection*. In the simple receiver shown, detection is accomplished with the help of a crystal detector. Dectection is the opposite of modulation, and is sometimes called demodulation. In the receiver shown, detection separates the amplitude modulation from the carrier.

All detectors are essentially rectifiers. The crystal detector, shown in the block diagram of the receiver on page 270, conducts effectively only in one direction. Thus, when it is connected across the selector tank circuit, the crystal acts as a rectifier. Its output is a pulsating direct current, pulsing at the RF rate but at the AF (modulation) amplitude. This is the first step in detection.

The process of detection is completed by filtering the pulsating DC. In the simple circuit shown below, the coils in the headset form part of the filter. They offer little impedance to AF but high impedance to RF. A capacitor across the headset completes the filter. It offers little impedance to RF but high impedance to AF. Thus the capacitor bypasses the RF around the headset while the AF signal goes through the headset.



Detector Output Filtering

Reproduction

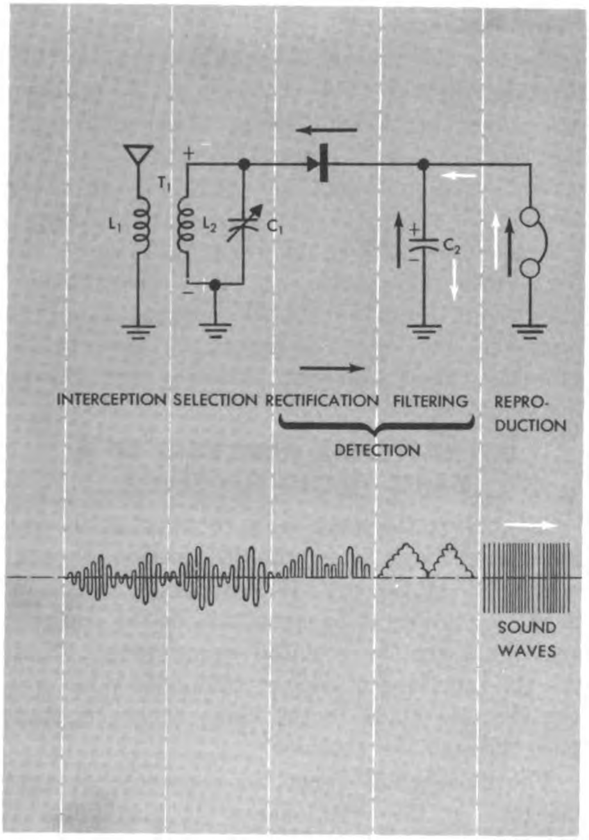
The current flowing through the coils of the headset varies at an audio rate. If the current can cause vibrations in the air at the same rate, the vibrations will constitute sound. For the reproduction of sound, the energy of the magnetic field which surrounds the coils is used. The strength of the magnetic field varies as the current varies. In turn, it causes a metal diaphragm to move back and forth at the same rate. This movement of the diaphragm sets up vibrations in the air and produces corresponding sound waves.

OPERATIONAL ANALYSIS OF A BASIC RADIO RECEIVER

To analyze the basic radio receiver, study the schematic diagram in the illustration on the next page. All the steps just discussed are shown here. The points to be considered in the analysis are: What are the essential components? What are the functions of each component? What are the changes made in the radio waves as they pass through the receiver?

The illustration shows the conventional connections of the components. It also gives a graphic representation of the wave as it passes through the receiver. When the radio wave passes or cuts across the antenna, RF voltage is induced across the antenna, and consequently across the primary of T1. As the current flows in L_1, the primary of T1, a voltage is induced in the secondary L_2. The antenna transformer usually has an air core and a step-up ratio which gives a small amount of signal gain.

The secondary of T1 is part of a tuned resonant circuit composed of L_2 and C_1. This tuned circuit offers maximum impedance to the frequency to which it is tuned. At this frequency, the voltage developed between the top of the tuned circuit and ground is maximum, and the voltage applied to the crystal is maximum. The tuned circuit offers little impedance to frequencies off resonance. For the off-resonant frequencies, only small voltages are developed across the tuned circuit and applied to the crystal. This means that the tuned circuit effectively shunts to ground the off-resonant frequencies and applies only the resonant frequency to the crystal. The waveform is still the modulated RF, with its amplitude increased by the step-up transformer.



Analyzing the Basic Tunable Receiver

The polarity at the extremities of the transformer changes with changes in the incoming signal. When the top of the tank circuit is positive, current flows through the circuit as indicated by the solid black arrows. When the top of the tank is negative, the crystal does not conduct, and the capacitor discharges as shown by the white arrows. Thus, the waveform is rectified, and the signal is changed to a pulsating DC. Note that the pulses still bear the imprint of the modulation after filtering, but this does not affect the intelligence.

Part of the filter is C_2, a small capacitor, usually about 250 mmf. It filters the RF component of the pulsating DC. The capacitor charges during the half cycles that the crystal conducts and discharges partially when the crystal is not conducting. In its charging and discharging, it absorbs almost all the swing of the RF pulses and follows the pattern of the modulation envelope, as the waveform shows.

A headset changes current pulses to sound waves by application of the electromagnetic principle. Each receiver contains two coils, a U-shaped permanent magnet which serves as the core of the coils, and a flexible diaphragm. As current travels through the coils, a changing magnetic field is developed. The strength of this field determines how much the diaphragm is attracted. When the field strength decreases, the diaphragm pulls away. The mechanical vibrations of the diaphragm cause the sound which is the intelligence.

LIMITATIONS OF A BASIC RADIO RECEIVER

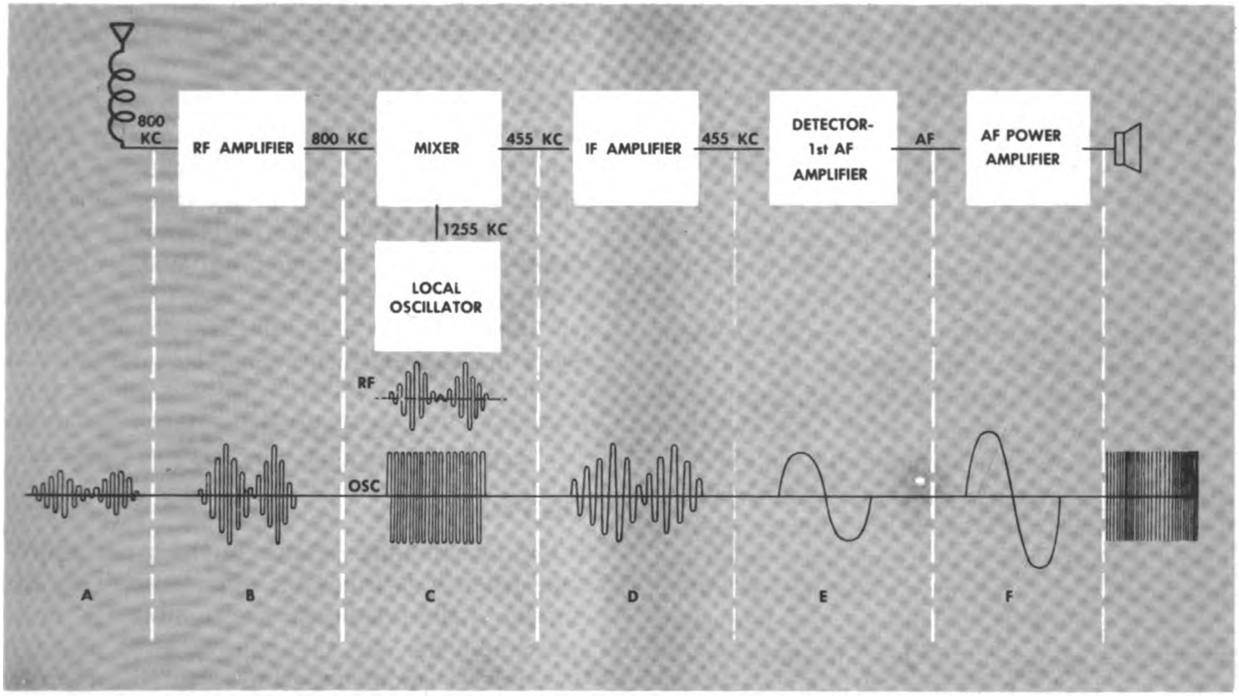
The basic receiver just described has its limitations. There is only one tuned circuit to improve the selectivity of the antenna. The slight step-up of the antenna transformer does little to improve the sensitivity of the receiver. This limits the effectiveness of the receiver to strong signals from nearby stations. On other signals, the output of the detector is not strong enough to develop magnetic fields around the headset coils to produce diaphragm vibrations.

An effective receiver needs more than these basic requirements. It must include circuits and stages to improve performance in terms of selectivity and sensitivity. It needs additional tuned circuits for selectivity and RF amplifier stages to improve sensitivity. It also needs AF amplifiers to increase the audio output so that a speaker can be used in place of the headset to reproduce the intelligence.

PRINCIPLES OF A SUPERHETERODYNE RECEIVER

The superheterodyne receiver is the type of receiver in most common use today. Almost all broadcast and communications receivers are now superheterodynes. Most receivers designed for radar frequencies are superheterodynes.

The superheterodyne receiver takes advantage of two important facts of radio operation. First, amplification of a radio signal at a low frequency can be more successful than amplification at a high frequency. Second, amplification of a signal by fixed tuned circuits is more successful than amplification by variable tuned circuits.



Block Diagram of a Typical Superheterodyne Receiver

The superheterodyne converts all input frequencies to a single, fixed lower frequency which is amplified by fixed tuned circuits. This process distinguishes the superheterodyne from other receivers in performance, for it provides outstanding sensitivity and selectivity over the whole tuning range. A block diagram of a superheterodyne receiver is shown in the illustration above.

Note the IF (intermediate frequency) amplifier stage. It is this stage which employs the fixed tuned circuits. The fixed frequency which it amplifies is called the intermediate frequency, because it is lower than any input radio frequency within the receiver's tuning range, but higher than audio frequencies. The receiver shown in the block diagram is a broadcast receiver with an IF of 455 kc. This frequency is substantially lower than the lower limit of the broadcast band which is 550 kc.

Note the local oscillator stage. The output frequency of this stage is combined with the output frequency of the RF amplifier stage to produce a new frequency — the intermediate frequency. This process is called heterodyning. The frequencies shown on the block diagram are the frequencies at which the various stages operate when the receiver is tuned to receive a signal of 800 kc. As you can see, the oscillator frequency in the block diagram is 1255 kc and the RF amplifier frequency is 800 kc. This means that the oscillator operates at a frequency 455 kc above the RF amplifier frequency.

The difference (or intermediate) frequency appears in the output of the mixer. The oscillator frequency is heterodyned against the RF amplifier frequency in the mixer stage. The output of the mixer contains the oscillator frequency, the RF amplifier frequency, the sum of these two frequencies, and the difference between these two frequencies. The fixed tuned circuits of the IF amplifier are tuned to 455 kc. Therefore, they accept the difference frequency but reject all the others.

The heterodyning process is illustrated by the waveforms shown in the block diagram. The RF modulated wave intercepted by the antenna is shown at A. The waveform at B, shows the same modulated carrier after it has been amplified by the RF amplifier stage. The waveform C, shows the output of the local oscillator. It is unmodulated and its amplitude is considerably higher than the amplitude of the waveform put out by the RF amplifier. The waveform at D, represents the amplified difference frequency. Note that it has the same modulation pattern as the RF carrier. The waveform at E, represents the detected audio frequency and the waveform at F, represents the amplified audio frequency.

Of course, the superheterodyne receiver is not confined to an input frequency of 800 kc. The RF amplifier tuned circuit is variable and can select any frequency in the broadcast band. The tuned circuit of the oscillator is also variable. It is ganged with the RF amplifier tuned circuit so that it is always 455 kc above the frequency to which the RF amplifier is tuned. Thus, the difference frequency between the oscillator frequency and the RF input frequency is always 455 kc. This arrangement for keeping frequencies separated by a fixed amount is called *tracking.* Tracking makes it possible for the frequency presented to the IF amplifier to always be the same, no matter what the RF input frequency may be. For low, broadcast, and medium frequencies, the oscillator usually tracks above the signal. For VHF and UHF, the oscillator usually tracks below the signal.

There is one major disadvantage to superheterodyne operation. If a local oscillator frequency of 1255 kc can mix with an input frequency of 800 kc to produce a difference frequency of 455 kc, this same local oscillator frequency of 1255 kc can also mix with an input frequency of 1710 kc to produce a difference frequency of 455 kc. Thus, the mixer section might present to the IF amplifier the signal from two different stations at the same time, both converted to the same IF. The IF amplifier would accept and amplify both at the same time. The demodulator would detect the signal of both at the same time. The intelligence of both would be present in the speaker at the same time. Such a mixture of signals would be confusing if not unintelligible.

This second signal which might interfere with the desired signal is called the *image frequency*. Image frequencies can best be prevented by selective tuning of the RF amplifier section. Highly selective RF amplifier tuned circuits, when tuned 455 kc below the oscillator frequency, will reject a frequency 455 kc above the oscillator frequency. In other words, the RF stage which is tuned to a frequency of 800 kc rejects the image frequency of 1710 kc.

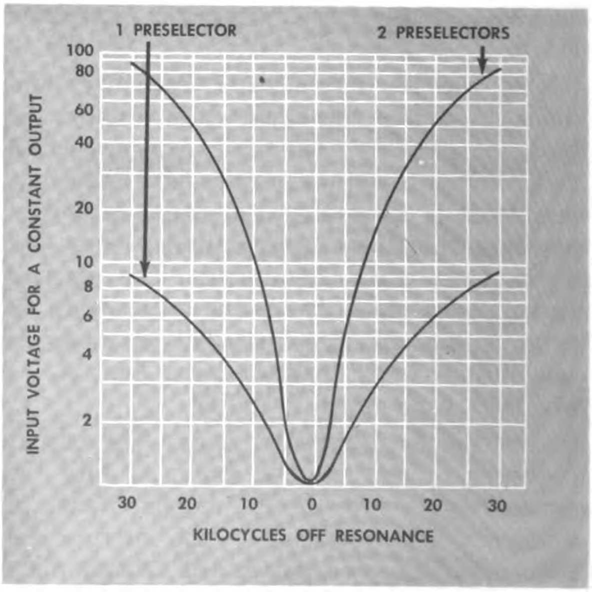
The superheterodyne receiver has uniform selectivity and sensitivity over its tuning range, because most of its gain is obtained in fixed-tuned circuits. These low frequency fixed-tuned amplifiers are more easily designed for high gain and selectivity. While the selectivity of the IF stages determines the overall adjacent channel selectivity of the superheterodyne receiver, an RF preselector having sufficient selectivity to reject image frequencies must be added.

ANALYSIS OF THE SUPERHETERODYNE RECEIVER

The Preselector

The RF amplifier section of the superheterodyne receiver is called the *preselector*. The RF circuits of the preselector are tunable RF voltage amplifiers. In the superheterodyne, the selectivity is more important than the gain of the RF amplifier stages since it is the selectivity which prevents the appearance of image frequencies.

One, two, or three RF amplifier stages may be used for the preselector section. Two is the usual number in a communications receiver, whereas, one is used in most house radios. The



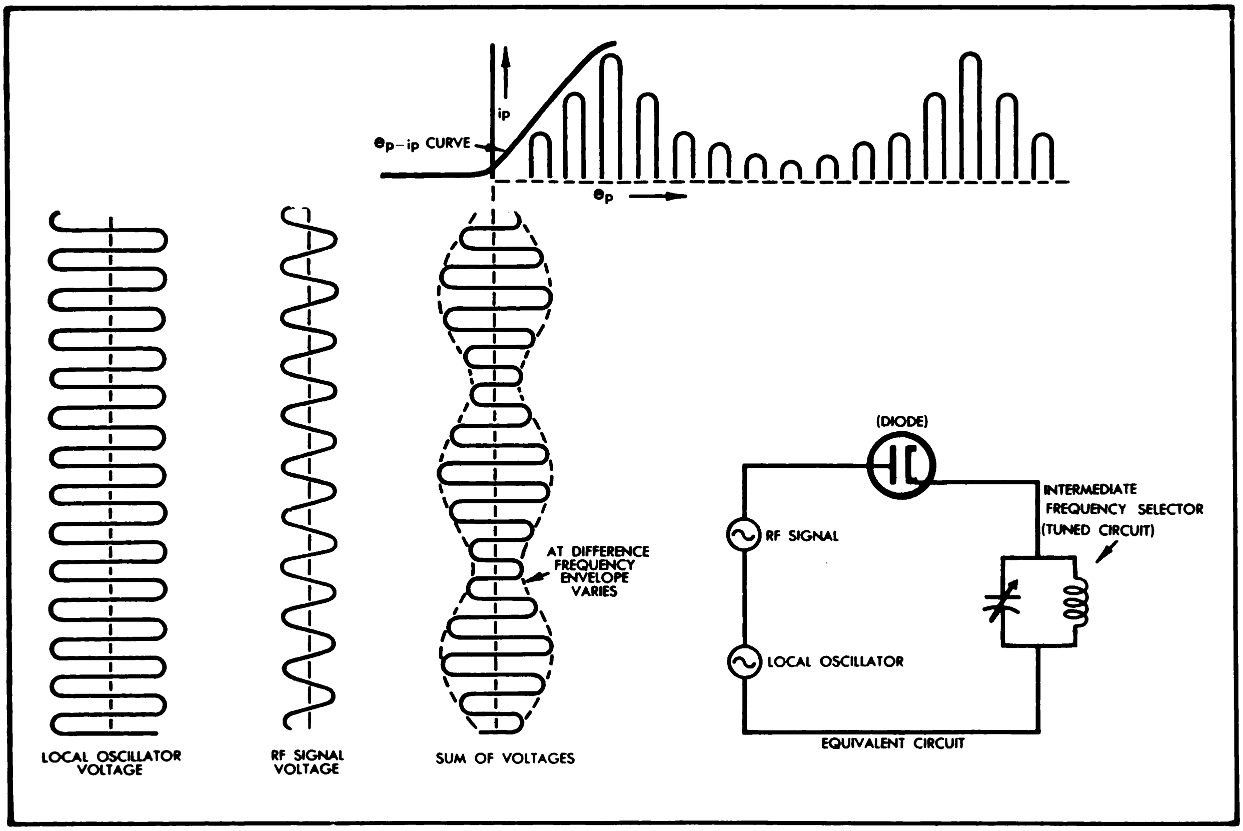
Selectivity of a Superheterodyne Preselector selectivity graphs on the preceding page show the amount of selectivity obtained with a preselector section of one and with a preselector of two RF amplifier stages. The vertical axis shows the relative input voltage required for a constant output. The horizontal axis shows the number of kilocycles off resonance. Notice that, with a preselector of one RF amplifier, a signal 30 kc off resonance must have about nine times the voltage of a signal at resonance to give the same response. With a preselector of two RF stages, a signal 30 kc off resonance must be about eighty times as strong as a signal at resonance to give the same response.

Frequency Conversion

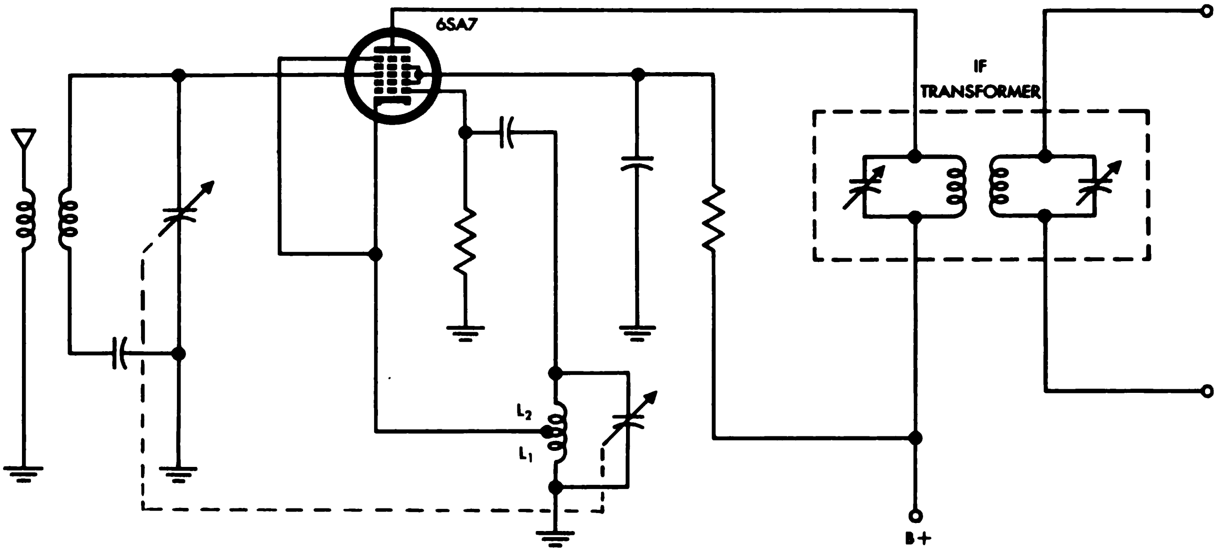
The heterodyning process that takes place in the superheterodyne receiver is called *frequency conversion.* The stage in which this frequency conversion takes place is called the mixer or converter, depending upon the arrangement of the circuit. When one multielement tube is used as a combination oscillator and frequency converter, the stage is called a *converter.* When one tube is used as the oscillator, and the output of the oscillator and signal frequencies are mixed in another tube, the tube in which the frequencies are combined is called a *mixer.*

Converters and mixers are also known as first detectors. The reason for this is that either a converter or a mixer tube must be nonlinear before its output will contain the difference, or intermediate, frequency. In other words, the tube must conduct more during the positive portions of the applied frequencies than during the negative portions — otherwise, the intermediate frequency would not be produced.

Frequency conversion is accomplished by mixing the RF frequency with the local oscillator frequency in a superheterodyne receiver. This mixing produces a frequency equal to the difference of the two frequencies. If the two frequencies are nearly the same, the difference frequency will be low. This frequency lies between the RF and audio frequencies, and is



Mixer Action



Pentagrid Converter

called the intermediate frequency. If the RF frequency is below 1000 mcps, it is usually amplified by an RF amplifier before being fed to the mixer. If higher, the RF signal is introduced directly into the mixer. After mixing, the IF frequency is amplified by an IF amplifier and then detected in the same manner as any other signal.

When two slightly different frequencies are added as shown in the illustration on page 275, the waveshape representing the combining of the local oscillator frequency and the RF signal frequency contains only these two frequencies. After distortion by the mixer the waveshape contains not only these two frequencies, but a frequency equal to the sum of the two frequencies and another equal to the difference between the two frequencies. Since it is the difference frequency that is desired, the tuned circuits in the IF amplifier select this frequency and amplify it. If the RF signal is modulated, the difference-frequency will also be modulated in the same manner.

Local Oscillator

The local oscillator must meet exacting requirements in frequency coverage, frequency stability, constant output, and correct tracking. The local oscillator may use any of the fundamental oscillator circuits. The modified Hartley is commonly used. Most VHF and UHF receivers use crystal-controlled local oscillators. To maintain frequency stability, the plate voltage of the oscillator is often regulated.

Pentagrid Converter

There are various methods of accomplishing the heterodyne action. It can be done, as shown in the illustration of a typical superheterodyne, by a separate oscillator tube feeding its output into the mixer tube; or it can be accomplished by a single tube (a converter) acting as an oscillator and mixer. At higher frequencies the local oscillator generally is a separate tube. Most house radios however, use the converter method in which a pentagrid tube acts as oscillator and mixer as shown in the illustration above.

The oscillator section of the tube is composed of cathode, first control grid, and the combined screen grids (as the anode). Feedback to maintain oscillation is provided by the autotransformer action of tapped coil L1-L2. Current flowing to the cathode through L1 induces the feedback in L2. The oscillator voltage appearing on the first control grid affects the flow of current through the tube. The RF signal voltage is applied to the second control grid.

Notice that in the frequency conversion circuit diagram shown the variable capacitor of the oscillator tuned circuit is ganged with a variable capacitor of the RF tuned circuit. This makes the oscillator tuned circuit track with the RF tuned circuit, and produce the correct IF. If the IF is 456 kc, the condensers are ganged so the local oscillator is always 456 kc above the frequency the RF amplifier is tuned to.

IF Amplifier

An IF amplifier, as shown in the illustration below, is basically an RF amplifier with a fixed tuned input and output. The tuned circuits act as bandpass filters, accepting the IF but rejecting other frequencies. Since they are fixed tuned, the tuned circuits have constant Q. With variable tuned circuits, Q varies with frequency. Response may be too sharp at one end of the frequency band and too broad at the other. In a superheterodyne, the response is more uniform because the IF circuits are fixed tuned. In addition, these fixed tuned circuits provide high selectivity and sensitivity.

Most IF transformers are double tuned with both a tuned primary and a tuned secondary. When less selectivity is desired, the primary is untuned. For greater selectivity, a third tuned circuit is sometimes inserted between primary and secondary.

The IF transformers used in any particular set may all be identical or they may vary from stage to stage depending on the design of the set. In one stage, for example, a higher or lower L/C ratio may be required in order to match impedances. Coils are shielded to prevent stray coupling.

The tuned circuits are fixed tuned by screwdriver adjustment. The adjustable component may be a capacitor or a coil. Most receivers designed for military use have an adjustable core in the coil. This type of tuning is called permeability tuning. With this type of tuning, the components are more compact, and there are no open capacitor plates to vibrate, corrode, or gather moisture.

The number of such stages used in a superheterodyne receiver depends on the operating frequency and on the amount of gain desired. At higher frequencies, sensitivity and selectivity decrease and more stages are required. Receivers for the UHF and VHF and radar may have many IF stages.

The Detector

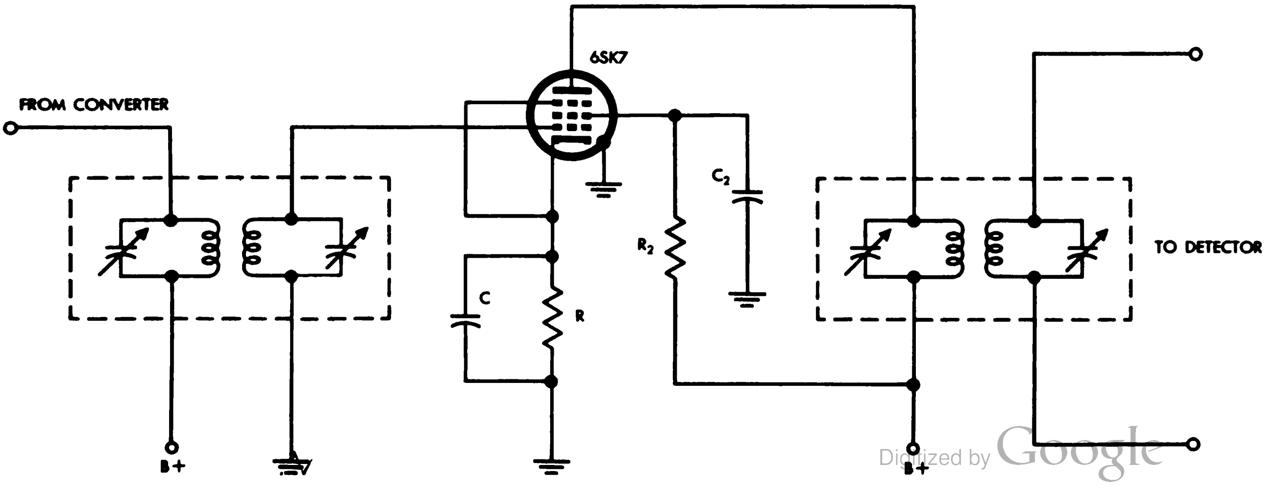
The output of the IF amplifier is fed into the detector. As was stated previously, the detector separates the intelligence from the modulated IF signal.

The detection of amplitude-modulated signals requires a nonlinear electrical network. An ideal nonlinear curve for detection is one that affects the positive half-cycles of the modulated wave differently than the negative half-cycles. It also distorts a wave with an applied voltage of zero average value so that the average resultant current varies as the signal frequency. The curve shown at A, in the illustration at the bottom of page 278, is called an ideal curve because it is linear on each side of the operating point P and does not introduce harmonic frequencies.

When the input to an ideal nonlinear curve is a carrier and its sidebands, the output contains the following frequencies:

1. The carrier frequency.
2. The upper sideband.
3. The lower sideband.
4. A frequency equal to the carrier minus the lower sideband (or, a frequency equal to the

IF Amplifier Stage



upper sideband minus the carrier), which is the original *signal frequency*.

5. A frequency of zero or a DC voltage.

The detector reproduces the signal frequency by producing a distortion of a desirable kind in its output circuit. When the output voltage of the detector is impressed upon a low-pass filter, which suppresses the radio frequencies, only the original or signal frequency is left.

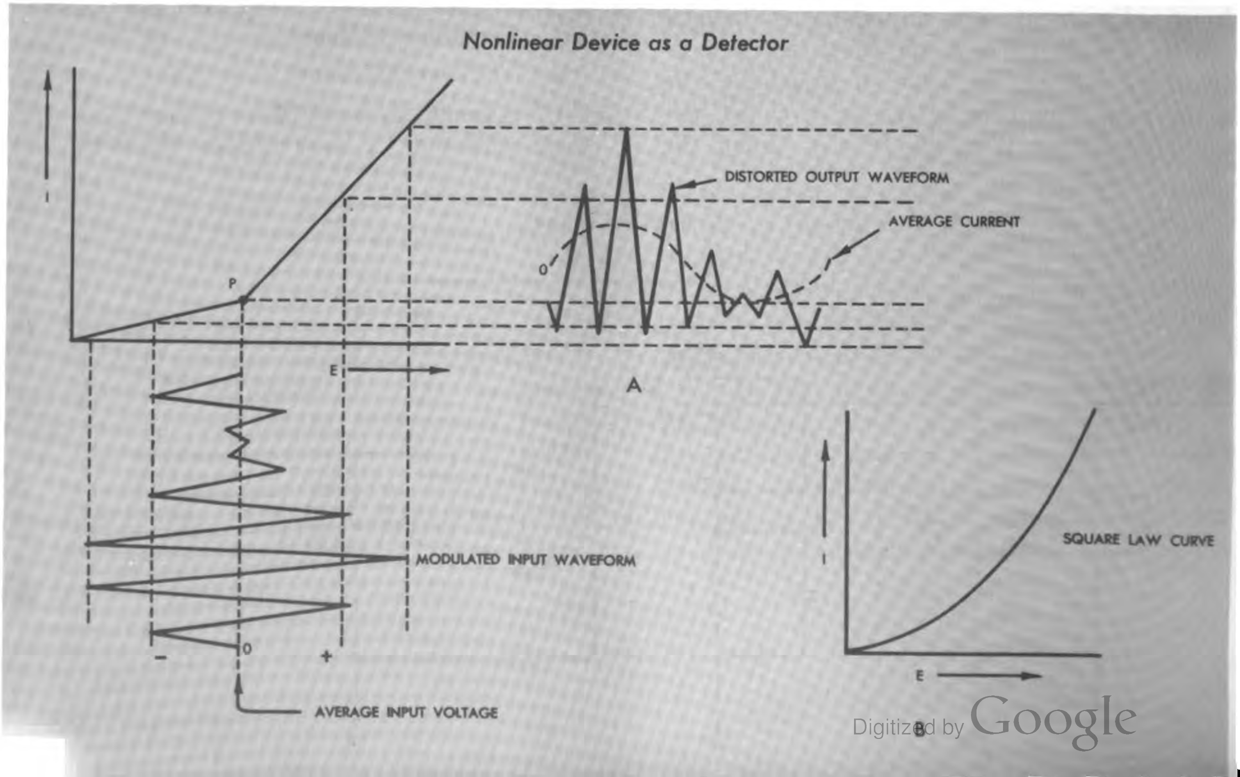
In some practical detector circuits, the nearest approach to the ideal curve is a square-law curve such as the one at B. The output of a device using this curve contains, in addition to all the frequencies which were listed, the harmonics of each of the input frequencies. The harmonics occur because voltage inputs which have a large amplitude are distorted differently than voltage inputs which have a low amplitude. The harmonics of radio frequencies can be filtered out, but harmonics of the signal frequencies, even though they produce an undesirable distortion, have to be tolerated. However, the square-law curve offers an advantage in that the output amplitude varies as the square of the input. In contrast, a linear curve produces an output which varies directly with the input. Therefore a square-law device produces a greater output for a given input.

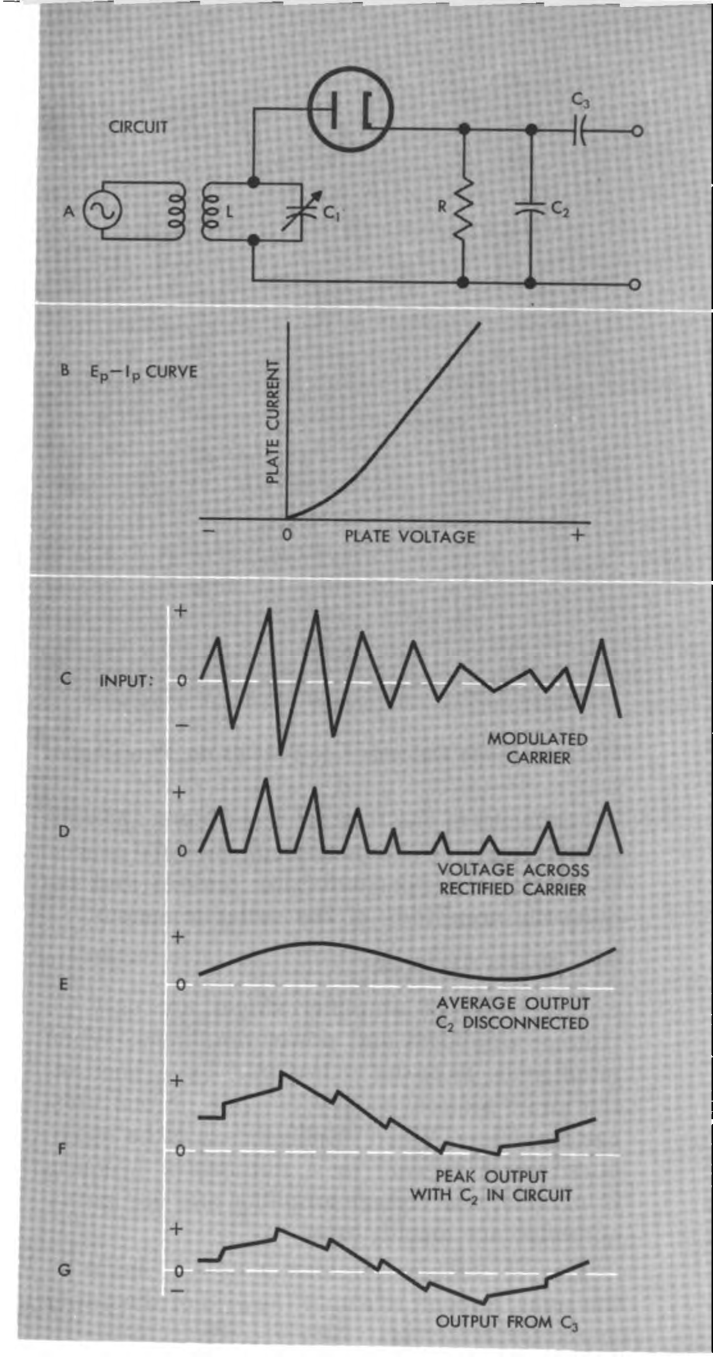
Diode Detector

One of the simplest and most effective types of detectors and one with nearly an ideal nonlinear resistance characteristic is the diode detector. Notice the E_p I_p curve at B, shown in the illustration to the right. This is the type of curve on which the diode detector at A operates. The curved part of the curve is the region of low plate current, and indicates that for small signals the output of the detector will follow the square law. For input signals with large amplitudes, however, the output is essentially linear in the positive direction from the operating point. This type of detector is classed as a power detector since it handles large input amplitudes without much distortion.

In the diode detector circuit on page 279, note that the modulated carrier is introduced into the tuned circuit LC_1. This circuit is designed so that the receiver has high selectivity. The waveshape of the input to the diode plate is shown at C. As a diode conducts only during positive half-cycles, it removes all the negative half-cycles and gives the result shown at D. The average output is shown at E. Although the average input voltage is zero, the average output voltage across R always varies above zero and has an average voltage of $\frac{1}{2} \times .637 \times$peak voltage for any positive half-cycle.

The low-pass filter, made up of the condenser C2 and the resistor R removes the RF (carrier) frequency, which, so far as the receiver is concerned, serves no useful purpose. The resultant output with C_2 in the circuit is a varying voltage which follows the peak variation of the modulated carrier (see F). The DC component produced by the detector circuit is still in the


Nonlinear Device as a Detector



Diode Detector

waveshape but may be removed by the condenser C_3, producing the AC voltage waveshape at G. In communications receivers, the DC component is often used to provide automatic volume (gain) control.

Diode detectors have the following advantages:

1. Ability to handle relatively high-power signals. There is no practical limit to the amplitude of the input signal.
2. Low distortion. Distortion decreases as the amplitude increases.
3. High efficiency. When properly designed, 90% efficiency is obtainable.
4. Ability to develop a readily usable DC voltage for the automatic gain control circuits.

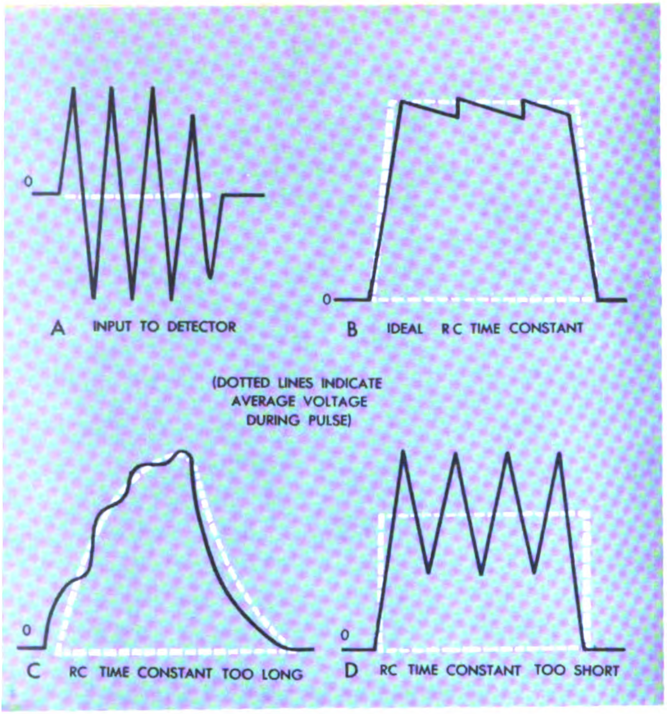
The diode detector has the following disadvantages:

1. Power is absorbed from the tuned circuit by the circuit. This reduces the Q and selectivity of the tuned input circuit.
2. No amplification occurs in a diode detector circuit.

Component parts must be carefully selected to obtain optimum efficiency in diode detector circuits. One very important fact to consider is the values of R and C_2, particularly in the case of pulse modulation. When a carrier modulated by a square pulse such as shown at A, in the illustration on page 280, is applied to an ideal diode detector, the detector produces a waveshape such as shown at B. Notice that the amplitude of this wave is great, while that of the high-frequency carrier shown at A is small.

If the time constant of RC_2 is too long, several cycles are required to charge capacitor C, and the leading edge is sloped as shown at C, on the next page. (Time constants and their effects on square waves are discussed in chapter 30.) After the pulse passes by, the capacitor discharges slowly and the trailing edge is exponential rather than square as desired. If the time constant is too short, both the leading and trailing edges are steep, but the condenser discharges considerably between cycles, and thus reduces the average amplitude of the pulse and leaves a sizable component of the carrier frequency in the output as at E.

The selection of the time constant is therefore a compromise between an output that is large but distorted, and one that is undistorted but small and only partly filtered. To achieve this compromise, the value of the load resistor R must be large. Since the total input voltage divides across R and the internal resistance of the diode when the tube is conducting, a greater proportion of the input voltage will be in the output where it is desired. On the other hand, the value of load resistance must not be made so high that the condenser C_2 must be made small enough to approximate the size of the inter-



Effect of Diode Detector Components on Waveshapes

nal capacity of the tube. If this occurs, the condenser C_2 will discharge through the tube capacity during the nonconducting periods and reduce the amplitude of the detector output.

The Audio Section

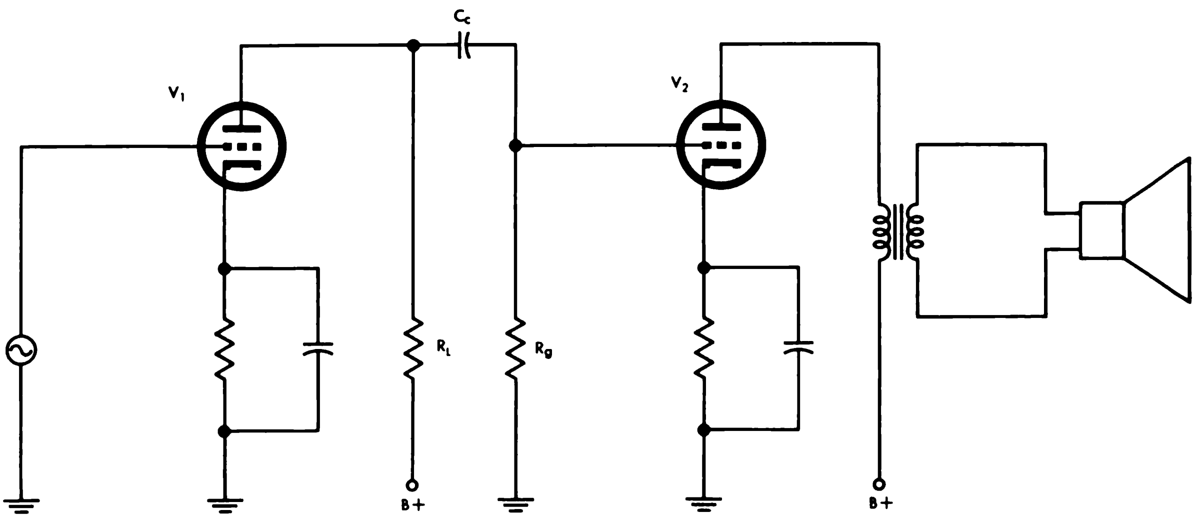
Two types of amplifier stages are used — voltage amplifiers and power amplifiers. When intelligence is to be reproduced by a headset, the audio section usually has one or two voltage amplifiers. When intelligence is to be reproduced by a speaker or a number of headsets in parallel, one voltage amplifier and one power amplifier are usually used.

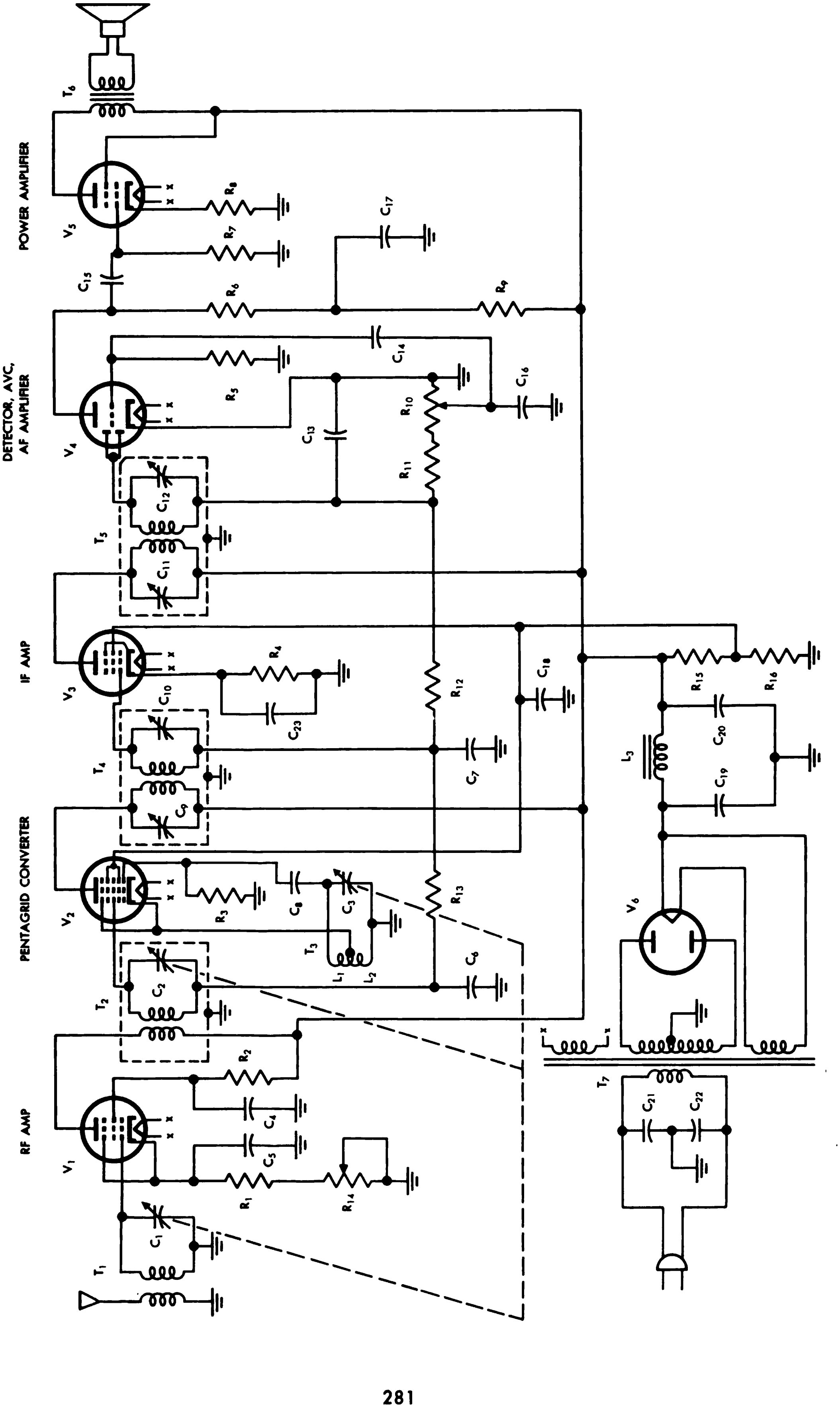
The AF voltage amplifier is used to build up the signal enough to excite the grid of the power amplifier or to energize the headset. Either a triode or pentode tube can be used as the voltage amplifier. In most cases, a medium-mu tube is used, especially when preceding RF amplifier stages provide sufficient gain. Resistance-capacitance coupling of the voltage amplifier to the power amplifier is conventional. The voltage gain for a medium-mu triode operating at frequencies around 1000 cps runs from 10 to 70. The load resistance of the voltage amplifier should be several times the plate resistance.

A power amplifier must efficiently convert DC power to AC power. If a triode is used, it should have low internal resistance. Maximum power transfer takes place when the load impedance equals the internal resistance. However, in practice, the load impedance is kept smaller than the plate resistance to prevent distortion. Power amplifiers generally use tetrodes or pentodes. These have higher power efficiency than triodes. Some tetrodes deliver nearly twice the output of a triode, even when the grid voltage of the tetrode is one-fourth that of the triode. For tetrodes and pentodes, a load impedance of about one-tenth of the plate resistance is used.

Study the voltage and power amplifier shown in the circuit diagram below. The generator in the grid circuit of V1, the voltage amplifier, represents the detector stage as a source of driving energy. The voltage output of V1 is capacitively coupled to V2, the power amplifier. The output of the power amplifier is transformer coupled to the speaker so that the high impedance of the power amplifier plate circuit can be matched to the low impedance of the speaker voice coil.

Audio Amplifier Section





Schematic of a Superheterodyne Receiver

THE COMPLETE SUPERHETERODYNE RECEIVER

To get an overall view of the superheterodyne receiver, examine the schematic diagram of the superhet receiver shown in the illustration on page 281. Although this receiver is somewhat simplified, it is representative of the receiver schematic you may see. Notice that it contains all of the circuits that we discussed: the RF amplifier (preselector), converter, IF amplifier, detector, audio section, and power supply. Each of these circuits will be briefly discussed.

The RF Amplifier. This stage is used to amplify the signal received from the antenna and to aid in discrimination against unwanted frequencies. It also aids in controlling the output volume of the receiver since it has a manual gain control in the cathode. R14 varies the cathode bias and therefore the gain of V_1.

The Converter. The pentagrid converter acts as an RF amplifier, mixer, and local oscillator. The action of the oscillator and mixer section produces the IF frequency. The secondary of T_2 is a tuned resonant circuit increasing the strength of the signal applied to the grid. Notice that T_1, T_2, and T_3 are tuned resonant circuits, tuned by condensers C_1, C_2, and C_3 which are ganged together. This is to insure that the resonant circuits are tuned simultaneously to the correct frequency and also to insure proper oscillator tracking.

The IF Amplifier. This stage is used to amplify the IF frequency. T_4 is tuned to the IF frequency.

The Detector, Audio Amplifier, and AVC. V_4 is a multipurpose tube. It is used as a detector and audio voltage amplifier and it also provides the automatic volume control (AVC) voltage. The purpose of this AVC voltage is to keep the output of the receiver relatively constant even though the signal received at the antenna may vary. The principle behind AVC is that as the incoming signal varies so does the audio voltage produced by the detector. Notice that the voltage developed by the detector causes the left side of R-11 to become negative. If the signal voltage increases, this end becomes more negative. This increase in negative voltage is then felt on the grids of V_2, and V_3 as an increase in bias, thus decreasing their gain. If the signal voltage decreases, the voltage developed across R-11 is less and the gain of V_2 and V_3 is increased.

Filters formed by C6-R13, C7-R12 eliminate the AC component from the AVC voltage fed back to the grids of the converter and IF amplifier.

The audio voltage developed across R10 is capacitively coupled to the grid of V4 where it is amplified by the triode section of this tube. The amplified signal is then capacitively coupled to the audio power amplifier V_5. R_8 provides bias for the power amplifier. Notice that R_8 is not bypassed by a capacitor. This provides degenerative feedback resulting in greater fidelity.

Resistor R10 is the manual volume control and is used to adjust the output level of the speaker. It functions by providing a larger or smaller portion of the audio signal developed across it to the grid of V_4.

The power supply is a conventional full wave rectifier circuit with a capacitor input filter. R15 and R16 form a voltage divider and bleeder circuit. Screen voltage for the converter and IF amplifier are taken from this divider. Capacitor C18 keeps the screens of both tubes at RF ground potential. Note the plate and screen voltages applied to the power amplifier. B+ is applied directly to the screen but is applied through the primary of transformer T6 to reach the plate. Because of the voltage drop across the plate transformer, the screen voltage is slightly higher than the plate voltage.

The principle of the superheterodyne receiver is of great importance to the observer as nearly every receiver that you come in contact with is a superheterodyne receiver. Just a few of the superheterodyne receivers with which you are familiar are listed below:

Radio Compass
VHF Command Set
UHF Command Set
BC-348 Liaison Receiver
Loran Set (APN-9, ARN-70)
Radar Receiver (APS-23, APS-42)
Omni Range Receiver
IFF Receivers

Chapter 30

Transients and Time Constants

The transmission of properly timed pulses of radio-frequency energy is the basis of a relatively new system of communications called pulse modulation. The main advantages of this system are that numerous voice or code messages and still pictures (facsimile) can be transmitted simultaneously from one transmitter without any interference from each other or from thunderstorms and man-made static. This system minimizes the amount of equipment that is necessary at a communications center to handle the traffic.

Instead of transmitting the continuous audio voltage variations of the voice as a broadcast radio transmitter does, the pulse transmitter transmits a number of very short portions of each audio sine wave — enough so that the receiver at the other end can reproduce the true sound. During an interval when the pulse transmitter is not sending one of these short pulses from one message, it is sending a short pulse from another.

One of the first applications of pulse modulation was radar. The basic principles of radar are simple. A short pulse of radio-frequency energy is radiated from the transmitting antenna toward a distant object. A very small amount of this energy is reflected from the distant object and returns to the receiving antenna. The amount of time required for the energy to go to the object and back is measured by the receiving equipment and interpreted in terms of yards or miles. Since radio-frequency energy travels through space at a rate of approximately 162,000 nautical miles per second, it travels .162 nautical miles in one millionth of a second or one microsecond. For example, if it requires 62 microseconds for the energy to make the round trip, the object must be 5 miles away. These small periods of time are accurately measured by electronic circuits.

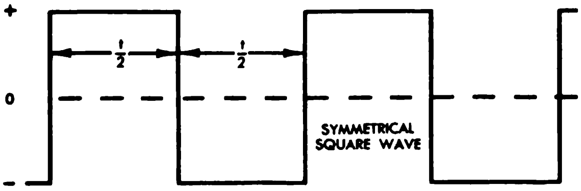
This chapter deals with some of the fundamental concepts behind pulsed signals — transients and time constants.

TRANSIENTS

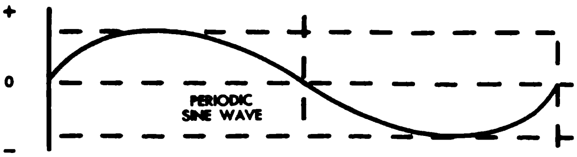
So far, you have been concerned with what goes on in a circuit when the circuit remains closed and current flows. One type of transient voltage is a voltage that is caused at the instant the circuit is closed or at the instant the circuit is opened. The voltage waveforms that you have studied in previous chapters have been sine waves or sinusoidal waves. A broader definition of a transient is any wave that is not a sine wave (nonsinusoidal). The most common of these are the square wave and the sawtooth wave. These waves may occur periodically at regular intervals like power-frequency sine waves, or they may occur at irregular intervals.

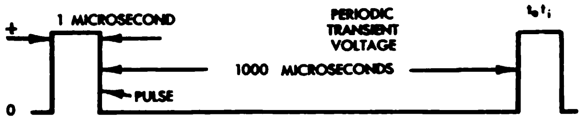
Square Wave

A square wave can be produced by periodically reversing the polarity of a direct voltage by means of a switch. A square wave voltage remains positive at a fixed value for a definite period of time and then becomes negative by the same amount for a like period of time. The number of times that the voltage changes from positive to negative per second is the frequency of the wave. The name of the wave is derived from the rectangular or square appearance of the positive and negative portions of the wave. It can be shown that a square wave consists of a fundamental sine-wave voltage whose frequency is equal to the frequency of the square wave, plus all the odd harmonics of this sine-wave voltage. (Assuming that the fundamental frequency of a voltage is 60 cps, the second harmonic has a frequency which is twice the frequency of the fundamental, or 120 cps. The third harmonic frequency is three times the frequency of the fundamental, or 180 cps, and so on.)



A Square Wave



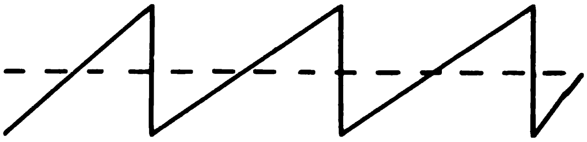


Short Rectangular Pulses

A square wave which consists of short pulses separated by a long time interval is very common in radar circuits. Such pulses are composed, like square waves, of a fundamental and harmonics to which a DC component is added. This type of square wave is called a rectangular pulse.

Sawtooth Wave

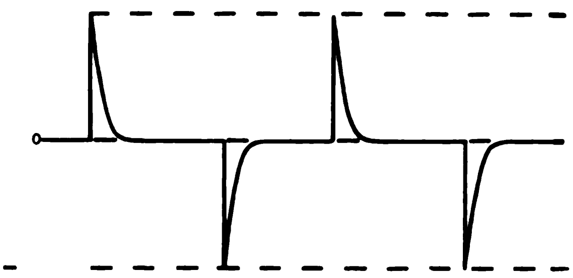
A sawtooth wave rises in a nearly linear manner from its minimum value for a definite period of time and then drops to its minimum value instantaneously. The number of times that the voltage rises and falls in this manner per second is the frequency of the wave. This type of wave is composed of the fundamental and all the odd and even harmonics of a sine-wave voltage whose fundamental frequency is the same as the frequency of the sawtooth wave. The name of the wave is derived from its sawtooth-like appearance.



Sawtooth Wave

Peaked Wave

A peaked wave is a wave which rises to a high positive peak and drops quickly; then falls to a negative peak and rises quickly. The sud-



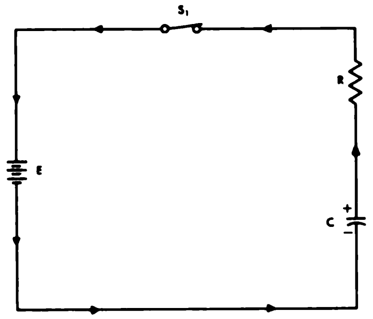
Peaked Wave

den rise and fall causes the sides of the wave to be very steep like those of the rectangular pulse. A peaked wave is composed of a fundamental sine wave and odd harmonics.

RC CIRCUITS

In your study of AC circuits, you have seen what happens when an AC sinusoidal voltage is applied across a resistor and capacitor in series. The output across the resistor, the output across the capacitor, and the input voltage all have the same shape — all are sine waves. Although the RC circuit causes a difference in the phase relationship between each of the outputs, the shape of the wave remains unchanged.

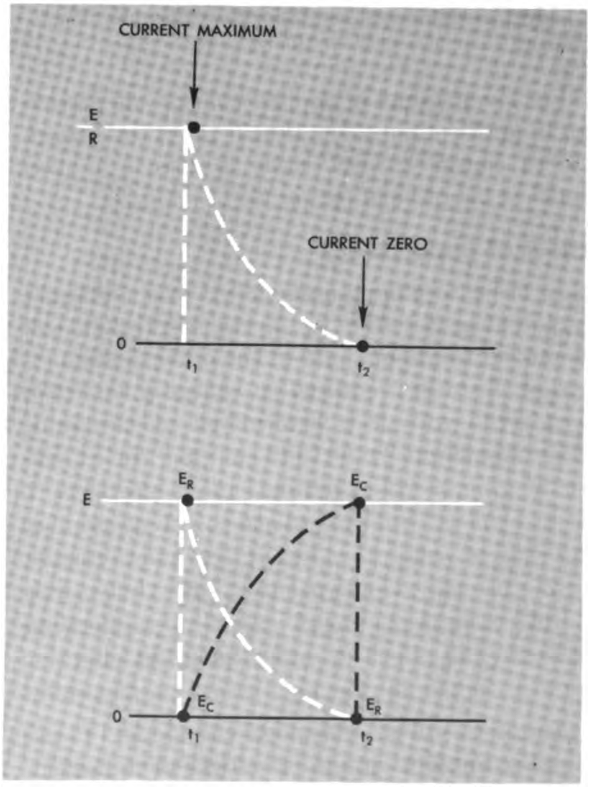
When a DC voltage is applied across an RC circuit, however, the output voltages across the resistor and across the capacitor have different waveshapes. Using the illustration at the upper right, you can analyze the outputs across the capacitor and resistor when a DC voltage is applied across an RC circuit. At the instant the switch is closed, the capacitor — which has no charge on it — acts like a short circuit. The resistor limits the flow of current to the value $\frac{E}{R}$. At this instant the entire voltage E is across the resistor and the voltage drop across the capacitor is zero. The current and the voltages at this instant are shown at time t_1 in the illustration of current and voltage waveshape.



Applying DC Voltage to an RC Circuit

If the switch remains closed, current flows in the direction indicated by the arrows, charging the capacitor with the polarity shown. The increasing charge on the capacitor offers an increasing opposition to the applied voltage and causes the current to decrease. After an interval of time, the capacitor is fully charged and current flow stops completely. At this instant (t_2) there is no voltage drop across the resistor; and the voltage across the capacitor is equal to the battery voltage E.

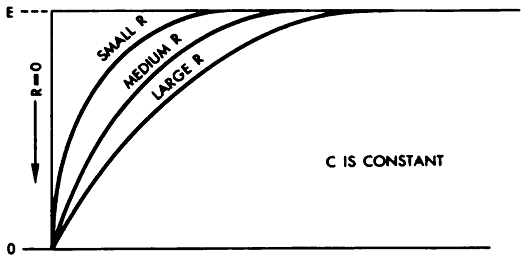
What factors in an RC circuit affect the time it takes for a capacitor to become fully charged and the current to become zero? An examination of the components of the circuit may give you the answer.



Current and Voltage in RC Circuit before and after Capacitor Charges

Effect of Resistance

The rate at which a capacitor charges depends upon the current flow in the circuit. With no resistance in the circuit, the capacitor would become charged instantaneously by a sudden surge of current. With resistance in the circuit, the current is limited by the value of the resistance. The smaller the value of the resistance, the greater is the current flow and the shorter is the time needed for the capacitor to charge to a given voltage. The larger the value of the resistance, the smaller is the current flow and the longer is the time needed for the capacitor to charge to a given voltage. The effect of resistance on charging time is shown in the illustration below.



Effect of Resistance on Capacitor Charging Time

For given values of applied voltage and capacitor in an RC circuit, the time it takes the capacitor to charge to any value is *directly proportional to the value of the resistor.*

Effect of Capacitance

In order to charge a capacitor to a difference of potential between its plates, electrons must be moved from one plate to the other. The number of electrons required to charge a capacitor to a given voltage is determined by the value of the capacitor. To be fully charged, the capacitor of greater value requires more electrons to be transferred from one plate to the other than the one of smaller value. If the same current flows through two capacitors, the time taken to fully charge the larger capacitor is therefore greater. The effect of capacitance on charging time is shown in the illustration to the right.

For given values of applied voltage and resistance in an RC circuit, the time it takes the capacitance to charge to any value is *directly proportional to the value of the capacitance.*

Effect of Applied Voltage

Suppose the applied voltage is increased and the values of R and C kept constant. At first glance you may think that the capacitor should take longer to fully charge to this higher potential difference. However, an increase in E causes a proportional current flow at every instant. The time for the capacitor to fully charge to the higher voltage therefore remains the same. Similarly, a decrease in the value of E has no effect on the time it takes a capacitor in an RC circuit to charge fully because the current decreases proportionally at each instant.

RC TIME CONSTANTS

Since the value of the resistor and capacitor affect the time it takes a capacitor in an RC circuit to charge, the values of R and C may be combined to indicate this time. The product of R (in ohms) and C (in farads) results in a product which is equal to time (in seconds).

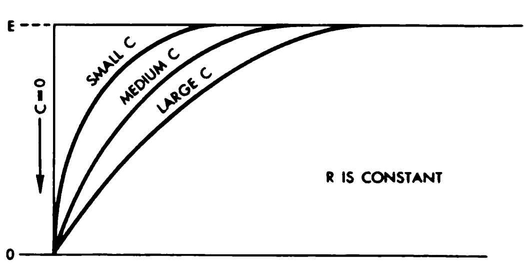
TC = R × C

What does this time (R×C) represent? It has been proved mathematically and verified by experimentation: If a DC voltage is applied to a series RC circuit for a time equal to the product of R and C, *the capacitor will have approximately 63% of the applied voltage across it.* The product of R and C is called a time constant (TC). Expressed mathematically,

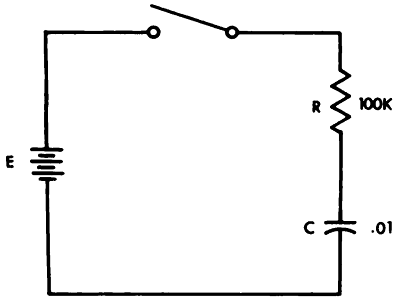
$$TC = R \times C$$

where TC is in seconds, R in ohms, and C in farads.

In discussing transients, you will find that TC is more useful when given in microseconds. Since the capacitance is usually given in microfarads, you can find the value of TC in microseconds by multiplying R in ohms by C in microfarads. The time constant of the RC circuit shown on page 287 is found as follows:



Effect of Capacitance on Capacitor Charging Time



A Typical RC Series Circuit

$$TC = R \times C$$
$$TC = 100{,}000 \times .01$$
$$TC = 1000 \text{ microseconds}$$

This means that 1000 microseconds after the switch is closed, the voltage (E_C) across the capacitor is 63% of 100 volts (63 volts). If R is changed to 10K, TC equals 100 microseconds, and E_C is 63 volts at the end of only 100 microseconds. If C is replaced by a 1-mfd. capacitor, 100,000 microseconds is required for E_C to reach 63 volts.

By Kirchoff's law,

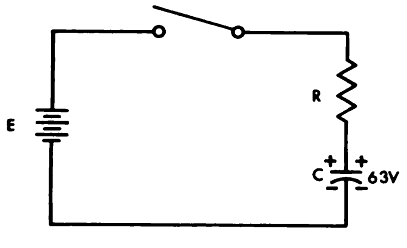
$$E_{app} = E_C + E_R$$
$$E_R = E_{app} - E_C$$
$$E_R = 100 - 63 = 37 \text{ volts}$$

Therefore, $E_R = 37$ volts at the end of one TC.

If the applied voltage in each of the preceding cases is changed to 10 volts, E_C is 6.3 volts and E_R is 3.7 volts at the end of one TC.

Plotting E_C and E_R

Suppose the switch in the RC circuit is opened after remaining closed for 1000 microseconds. Since the capacitor has no path through which to discharge, it retains its charge of 63 volts. The voltage in the circuit is as shown below. There is no voltage across the resistor because no current is flowing.



Voltage E_C with Switch Open

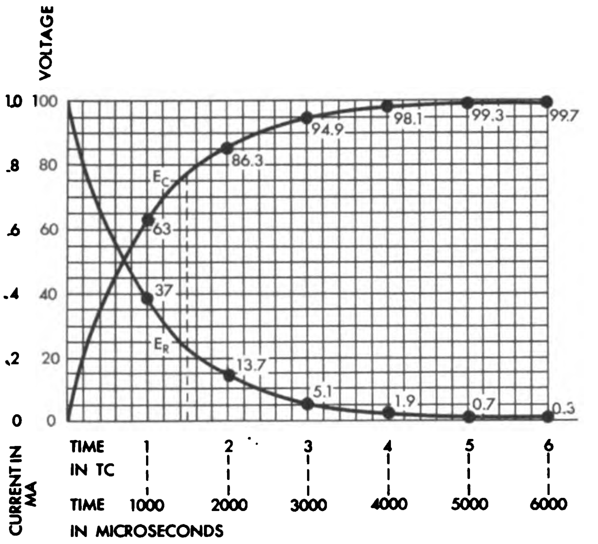
If the switch is closed after a short time, the capacitor begins to charge again. However, E_C is of such polarity as to oppose the applied voltage E. Therefore the voltage which causes current to flow is now equal to 37 volts. The capacitor therefore can charge only an additional 37 volts — to 100 volts. In 1000 microseconds from the time the switch is closed, the capacitor charges an additional 63% of 37 volts (23.3 volts). The total voltage across the capacitor is 86.3 volts and across the resistor 13.7 volts at the end of this time. If the switch had remained closed continuously for 2000 microseconds (2TC), the same result would have been obtained.

In similar fashion, you can see that the capacitor acquires an additional 63% of the difference in voltage between E and E_C in the third TC interval. Thus, 63% of 13.7 volts (8.6 volts) is added to the capacitor in this time interval. At the end of 3000 microseconds (3TC), $E_C = 94.9$ volts and $E_R = 5.1$ volts. The voltages E_C and E_R at the end of 4,000, 5,000, and 6,000 microseconds may be obtained in similar manner.

Voltages and Current in RC Circuit Shown				
Number of Micro-seconds	Number of TC	E_C (Volts)	E_R (Volts)	I (ma)
0	0	0	100	1.0
1000	1	63	37	.37
2000	2	86.3	13.7	.137
3000	3	94.9	5.1	.051
4000	4	98.1	1.9	.019
5000	5	99.3	.7	.007
6000	6	99.7	.3	.003

The table shows the voltages E_C and E_R at the end of each TC interval after the switch is closed in the circuit to the upper left. Check the voltages at the end of 4000, 5000, and 6000 microseconds by the same method of calculation used for the first three TC intervals.

Theoretically, a capacitor never becomes completely charged. However, for practical purposes, a capacitor is considered to be completely charged at the end of 5TC. Note that, at the end of this time, the voltage on the condenser is 99.3 volts or 99.3% of the applied voltage.



Curves E_R and E_C in RC Circuits

The last column shows the value of the current (I) at the end of each time constant. Since the voltage across the resistor results from the current flowing through it, and both E_R and R are known, you can find the current (I) in the series circuit from the relationship $I = \frac{E_R}{R}$. Note that the current and E_R drop in value at the same rate.

Plot the values shown in the table for E_C and E_R and connect these values with smooth curves. The result is shown in the illustration.

Note that E_C and E_R are reverse curves and that their sum at every instant is equal to 100 volts. By using the appropriate values for the coordinates, the current curve and the curve of voltage across the resistor are plotted as one curve.

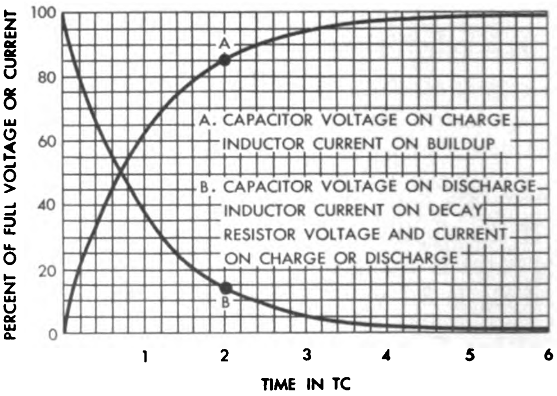
You can obtain the values of the current and both voltages directly from the graphs for any desired time after the switch is closed. For example: 1500 microseconds (1.5 TC) after the switch is closed, $E_C = 78$ volts, $E_R = 22$ volts, and $I = .22$ ma.

UNIVERSAL TIME CONSTANT CHART

Curves similar to those plotted for the preceding RC circuit may be plotted for each series RC circuit. However, one set of curves (Universal Time Constant Chart) may be conveniently used with any series RC circuit regardless of the values of E_{app}, R, and C. On these curves, the number of TC is plotted horizontally and the percentage of voltage or current is plotted vertically. You obtain the voltages E_C and E_R at any instant by multiplying the percentages obtained from the curves by the applied voltage. To obtain the value of current at any instant, multiply the percentage obtained from the curve by the current $\left(\frac{E}{R}\right)$ at the instant of closing the switch.

Using the Chart

You can see from the illustration below that the universal time constant curves may be used for purposes other than finding voltages and current as the capacitor charges. The capacitor-charging curve (curve A) may also be used for finding inductance current on buildup in an L-R series circuit. The current and resistor voltage curve (curve B) may also be used for finding the capacitor voltage on discharge, as well as the inductance current on decay in an L-R series circuit. All these uses will become more understandable as you proceed. The examples which follow employ the curves only in series RC circuits in which the capacitor is charging.



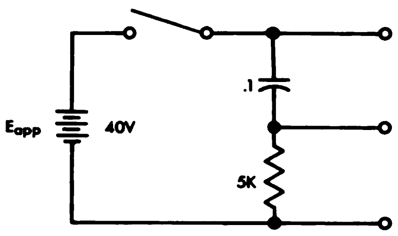
Universal Time Constant Chart

Example 1

Find E_C, E_R, and I,1200 microseconds after the switch in the series RC circuit shown is closed.

Solution:

$$TC = R \times C$$
$$TC = 5000 \times .1$$
$$TC = 500 \text{ microseconds}$$



Circuit for Example I

To find the number of TC in 1200 microseconds,

$$\text{Number of TC} = \frac{T}{TC} = \frac{T}{R \times C}$$

$$\text{Number of TC} = \frac{1200}{500} = 2.4$$

From curve A for 2.4 TC,

$$\begin{aligned} E_C &= 91\%\ E_{app} \\ &= (.91)\ (40) = 36.4 \text{ volts} \\ &= E_{app} - E_C \\ &= 40 - 36.4 = 3.6 \text{ volts} \end{aligned}$$

or using curve B

$$\begin{aligned} E_R &= 9\% \text{ of } E_{app} \\ &= (.09)\ (40) = 3.6 \text{ volts} \end{aligned}$$

$$I = \frac{E_R}{R}$$

$$I = \frac{3.6}{5000} = .72 \text{ ma}$$

Example 2

Find the time required for the capacitor in the circuit of example 1 to charge to 30 volts.

Solution:

$$\frac{E_C}{E_{app}} = \frac{30}{40} = 75\%$$

$$E_C = 75\%\ E_{app}$$

From curve A for 75%,

$$\text{Number of TC} = 1.4$$

$$\frac{T}{RC} = 1.4$$

$$T = (1.4)\ (500)$$

$$T = 700 \text{ microseconds}$$

Example 3

Find the time required for the current in the circuit of example 1 to drop to 2 ma.

Solution:

$$E_R = IR$$

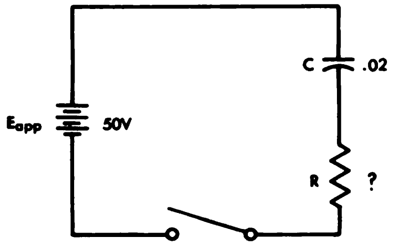
$$E_R = (.002)\ (5000)$$

$$E_R = 10 \text{ volts}$$

$$\frac{E_R}{E_{app}} = \frac{10}{40} = 25\%$$

From curve B for 25%, Number of TC = 1.4. Proceed as in example 2 to find T = 700 microseconds.

Example 4



What value of R is needed in the series RC circuit shown above for C to charge to 35 volts in 480 microseconds.

Solution:

$$\frac{E_C}{E_{app}} = \frac{35}{50} = 70\%$$

Using curve A, E_C charges to 70% E_{app} in 1.2 TC.

$$1.2 \text{ RC} = 480$$

$$1.2\ (R)\ (.02) = 480$$

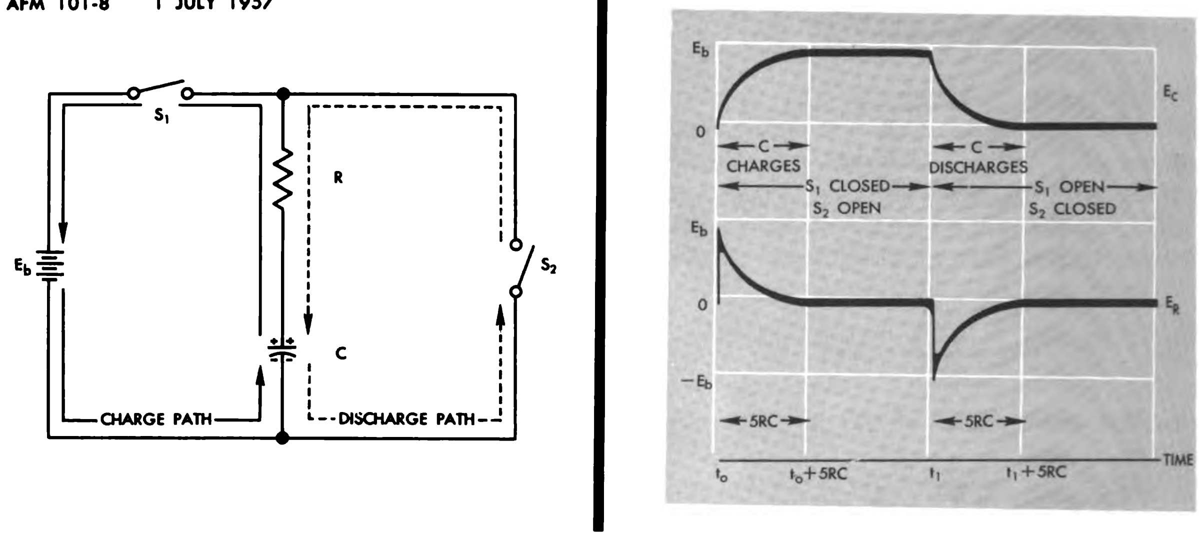
$$R = \frac{480}{.024} = 20000 \text{ ohms}$$

In similar manner, you can find the size of the capacitor if the value of the resistance is known.

Discharging the Capacitor

In the circuit, Charging and Discharging a Capacitor on page 290, assume that S_1 is closed and S_2 open. After an interval equal to five TC, the capacitor is fully charged, the current is zero, and E_R is zero. The charging of the capacitor is represented by the waveshapes shown in diagram B.

If S_1 is opened and S_2 closed after the capacitor is fully charged, the capacitor is provided with a path through which to discharge. Since the battery is no longer in the circuit, the only source of energy is that which has been stored in the electrostatic field of the capacitor. As the capacitor discharges, it causes current to flow through R. The voltages E_C and E_R are

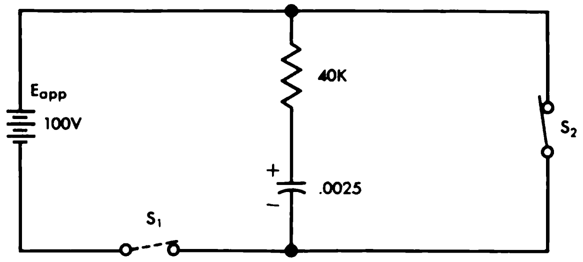


Charging and Discharging a Capacitor

equal at every instant. This fact is evident if you consider the capacitor a source of voltage and the resistor a load. By Kirchoff's law, E_C is equal to E_R and opposite in polarity. The arrows in diagram A show that the direction of current through R when the capacitor is charging is the reverse of its direction when the capacitor is discharging. In diagram B, this reversal of polarity is indicated by showing E_R below the zero line while the capacitor is discharging.

The rate at which the capacitor discharges, like the rate at which it charges, depends upon the time constant, RC. The voltage across the capacitor drops 63% of its voltage in the first TC and 63% of the remaining voltage in each succeeding TC. At the end of 5TC, the capacitor is fully discharged, the current flow is zero, and the voltage across the resistor is zero. The voltages E_R and E_C as the capacitor discharges are shown in diagram B. Note that E_R and E_C are equal at every instant but opposite in polarity. Curve B of the universal TC chart may be used for finding E_C and E_R as shown in the example below:

Example



In the circuit shown, S_1 is closed and S_2 open until the capacitor is fully charged. S_1 is then opened and S_2 is closed. Find E_C, E_R, and the current flow 130 microseconds after the capacitor has begun to discharge.

Solution:

$$TC = RC = (40000)\ (.0025)$$

$$TC = 100 \text{ microseconds}$$

$$\text{Number of } TC = \frac{T}{TC} = \frac{130}{100} = 1.3\ TC$$

From curve B for 1.3 TC,

$$E_C = 27\%\ E_{app} = 27\%\ (100)$$

$$E_C = 27 \text{ volts}$$

$$E_R = -27 \text{ volts}$$

$$I = \frac{E_R}{R}$$

$$I = \frac{27}{40000} = .675 \text{ ma}$$

SQUARE WAVE APPLIED TO RC CIRCUITS

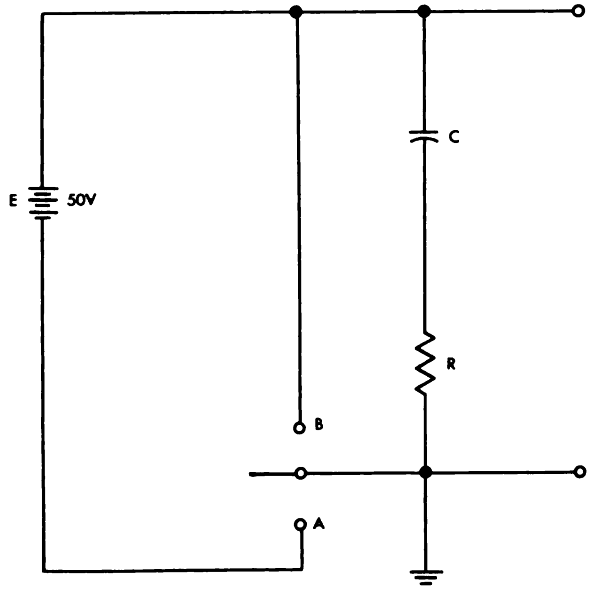
When a square wave of voltage is applied across a series RC circuit, the output waveshapes across the resistor and the capacitor are different. This was shown in the previous discussion. Note that the applied square wave of the voltage was distorted in these outputs. In some radar circuits this distortion is desired; in others it is not. The values of R and C and the width of the square wave affect the amount of distortion in the output.

Frequency and Period of a Square Wave

The voltage applied across the resistor and capacitor in series in the circuit shown below is a square wave of voltage. When the switch is in position A, the voltage across the RC circuit is 50 volts; when the switch is in position B, the voltage is zero. The square wave shown represents the alternate changing of the position of the switch, if the switch is kept in each position for the same length of time. Such a square wave is called a *symmetrical* square wave.

The time from t_0 to t_1 during which the switch is kept in position A is called an *alternation;* the time from t_1 to t_2 represents another alternation. The alternations that begin at times t_0, t_2, and t_4 are called positive-going alternations; those that begin at times t_1 and t_3 are called negative-going alternations. A *cycle* consists of one positive-going alternation and one negative-going alternation. The time of one cycle is called a *period* and is represented by the letter T. The frequency of a square wave and its period have the same reciprocal relationship as the frequency and period of a sine wave; that is, $T=\frac{1}{f}$ and $f=\frac{1}{T}$, where T is in seconds and f in cycles per second.

If the switch is placed alternately in position A and position B for 500 microseconds each, the time of each alternation (t) is 500 microseconds.



To find the frequency of this symmetrical square wave, proceed as follows:

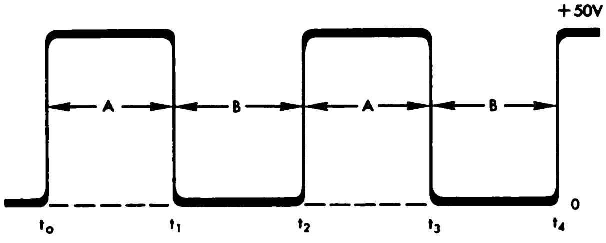
$$T = 2\,t$$
$$T = 1000 \text{ microseconds}$$
$$f = \frac{1}{T}$$
$$f = \frac{1}{.001}$$
$$f = 1000 \text{ cps}$$

The time of an alternation of a symmetrical square wave may be found if the frequency is known. For example, if the frequency is 500 cps, the time of an alternation is found as follows:

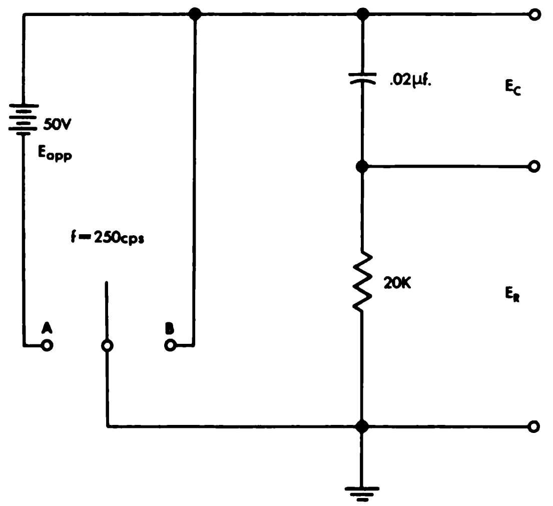
$$f = \frac{1}{T}$$
$$500 = \frac{1}{T}$$
$$T = \frac{1}{500} \text{ sec.}$$
$$T = \frac{1}{500} \times 10^6 \text{ microseconds}$$
$$T = 2000 \text{ microseconds}$$
$$t = \frac{T}{2}$$
$$t = 1000 \text{ microseconds}$$

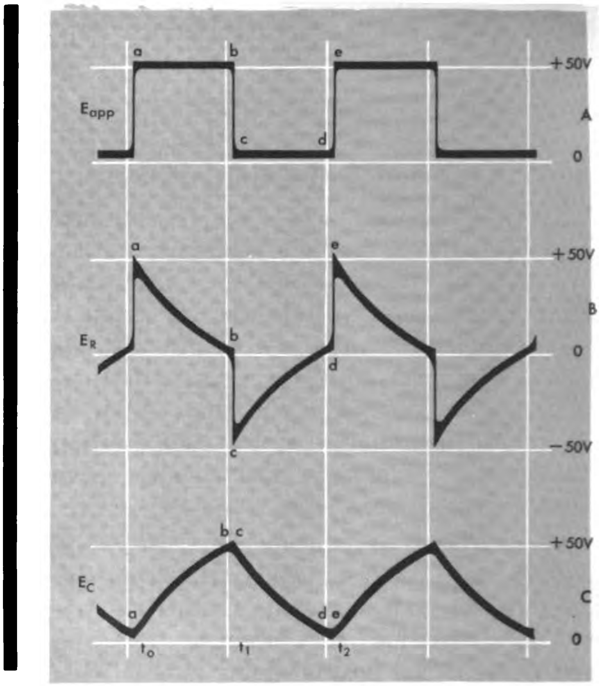
Applying a Square Wave

If the switch in the circuit shown on the next page is placed alternately in position A and position B for 2000 microseconds each, two cycles of the applied voltage may be represented by the square wave shown as E_{app}. The voltage across



Applying a Square Wave to an RC Circuit





Circuit and Output Waveshapes for t/RC = 5

the resistor (E_R) and across the capacitor (E_C) is found by using the universal time constant chart in the following manner.

WHEN $\frac{t}{RC} = 5$. At time t_0, the switch is placed in position A and the voltage applied to the circuit is 50 volts. At that instant, the entire voltage appears across the resistor and zero voltage across the capacitor as shown in diagrams B and C (point a). Since the switch is kept in position A for 2000 microseconds, 50 volts is applied for this length of time and the capacitor charges. At time t_1 — an instant before the switch is thrown to position B — the capacitor has charged to some voltage value and the voltage on the resistor has dropped accordingly.

To determine voltages E_C and E_R at the end of 2000 microseconds, find the number of time constants contained in 2000 microseconds as follows:

$$TC = R \times C$$

$$TC = 20000 \times .02$$

$$TC = 400 \text{ microseconds}$$

$$\frac{t}{TC} = \frac{2000}{400} = 5$$

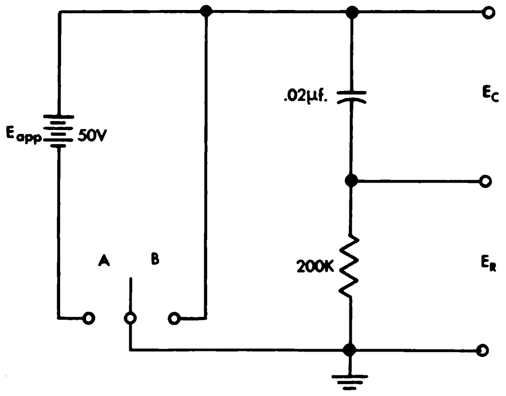
From the universal time constant chart, E_C is approximately equal to the applied voltage and E_R is approximately zero. This is shown as point b in diagrams B and C. If you connect points a and b in the diagrams by curves similar to those of the universal time constant chart, voltages E_R and E_C are obtained for 2000 microseconds. Note that $E_R + E_C = E_{app}$ at each instant.

At time t_1, the switch is placed in position B. This removes E_{app} and allows the capacitor to discharge. The three voltages are now represented by point c in the respective diagrams. The capacitor voltage remains at 50 volts and the voltage E_R is equal to E_C but opposite in polarity, because the current through it has reversed direction.

During the time t_1 to t_2, the capacitor discharges and both E_C and E_R decrease. At time t_2 the voltages E_C and E_R may be found by using curve B of the universal time constant chart. For five TC, voltages E_C and E_R are zero as shown at point d. Connect points c and d of E_R and E_C by curves similar to those of the universal TC chart. The waveshapes E_R and E_C shown from time t_0 to time t_2 result from one cycle of applied square wave. If the switch is thrown back to position A, a new cycle starts and the action of the first cycle is repeated.

These are the conditions that exist in the circuit only if the values are as shown; that is, C = .02 mfd., R = 20 K, and t = 2000 microseconds. If any one of the values is changed, $\frac{t}{RC}$ changes and the resultant waveshapes vary accordingly. In the circuit discussed, $\frac{t}{RC} = 5$; the capacitor charged completely on each positive alternation and discharged completely on each negative alternation.

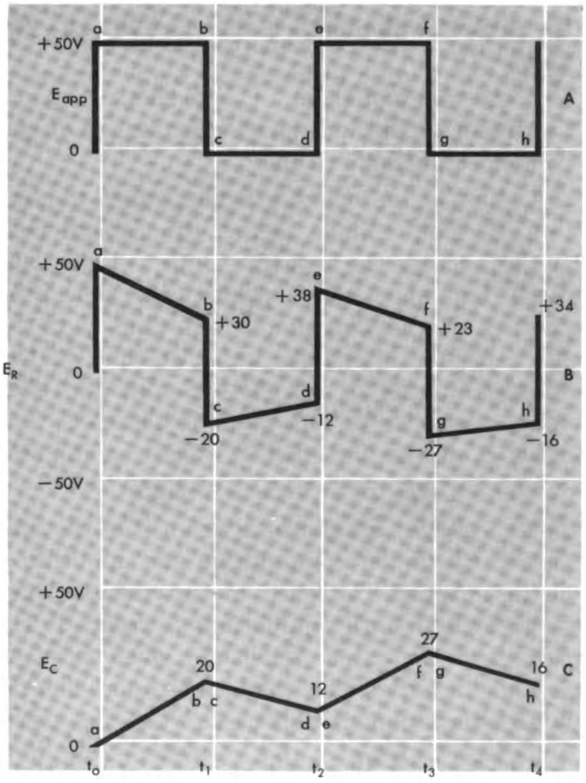
CHANGING THE RC CONSTANT. The effect of varying the value of RC on the output waveshapes may be shown by changing the value of R in the previous circuit. Replace the 20 K resistor by a 200 K resistor as shown in the illustration to the right. In this case, RC is 4000 microseconds and the value of $\frac{t}{RC}$ is 0.5.



Circuit for t/RC=0.5

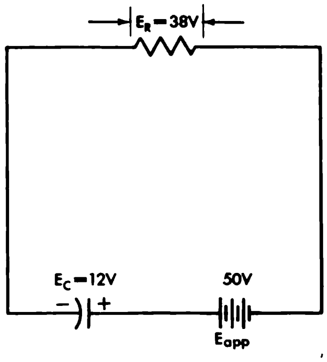
The waveshapes are obtained as before by using the universal time constant chart. These waveshapes are shown in the illustration to the right. At the instant (t_0) of closing the switch, the voltages E_{app}, E_R, and E_C are as shown at point a. Since the square wave is applied for 2000 microseconds, the capacitor charges and the voltage E_R drops. From the universal TC chart for 0.5 TC, E_C is approximately 40% E_{app} and E_R is approximately 60% E_{app}. As shown in the diagrams at point b, $E_C=20$ volts and $E_R=30$ volts. Note that, although the capacitor charges for the same length of time (2000 microseconds) as before, it charges at a slower rate because of the larger resistor and is charged to only 40% E_{app}.

At time t_1, the switch is placed in position B. The capacitor is now the source of voltage. Both E_C and E_R are equal but opposite in polarity. From time t_1 to t_2, the capacitor discharges. The voltages E_R and E_C at the end of 2000 microseconds may be found by using curve B of the universal TC chart. At the end of 0.5 TC, both voltages have dropped to 60% of 20 volts. As shown at point d in diagrams B and C, E_R -12 volts and E_C is 12 volts. Note that the sum of E_R and E_C is equal to zero — the applied voltage.



Waveshapes for t/RC=0.5

At time t_2, the switch is placed in position A again. At the instant of closing the switch, the capacitor has 12 volts across it. Since E_{app} is 50 volts and E_C is in opposition as shown in the illustration to the right, E_R is 38 volts. These voltages are shown at point e of the waveshape diagrams.



E_c Opposes E_{app}

As the capacitor charges from time t_2 to t_3, the voltage E_R drops. From curve B of the universal TC chart for 0.5 TC, you find that E_R drops to 60% of 38 volts or 22.8 volts. Note that 38 volts is used since the voltage across the resistor at time t_2 is $E_{app}-E_C$ as shown in the bottom illustration on page 293. The voltage E_C at time t_3 may be obtained by subtracting E_R from E_{app} as follows:

$$E_C = E_{app} - E_R$$

$$E_C = 50 - 23$$

$$E_C = 27 \text{ volts}$$

Another method of obtaining E_C is to use curve A of the universal TC chart. The capacitor has a 12-volt charge on it and charges toward 50 volt — a difference of 38 volts — in the time interval which is equivalent to 0.5 TC. From the curve, you find the capacitor changes 40% of 38 volts or 15.2 volts. The total voltage on the capacitor at time t_3 is therefore 27 (12+15) volts.

These values of E_R and E_C are shown at point f in the diagrams and the waveshapes are drawn by connecting points e and f.

At time t_3, the switch is again thrown to position B. The voltage across the capacitor is applied across the resistor. Both voltages are equal but opposite in polarity as shown at point g in the diagram. From time t_3 to t_4, the capacitor discharges until, at time t_4, the voltages E_C and E_R become +16 volts and −16 volts, respectively. This value is obtained from curve B of the universal TC chart as follows:

$$60\% \text{ of } 27 \text{ volts} = 16 \text{ volts}$$

These voltages are shown at point h in the diagrams. By similar computation and application of the unversal TC chart, the waveshapes may be computed until equilibrium is reached; E_R will have the same amplitude when positive as when negative. This would take approximately 15 alternations in this case. Note that when $\frac{t}{RC}$ is 5, as in the previous case, equilibrium is reached at the end of the first cycle. At equilibrium, the waveshape E_R is an AC wave which has equal amplitudes above and below the zero axis line.

RECTANGULAR PULSE APPLIED TO RC CIRCUITS

If the switch in the circuit shown below is placed in position A for 50 microseconds and in position B for 150 microseconds, the waveshape which results is not symmetrical. Since the time (t_1) of the positive-going alternation is less than the time (t_2) of the negative-going alternation, this type of square wave is called a *rectangular* wave or *pulse*. For this type of wave, the time of one cycle or period (T) is found by adding t_1 and t_2 as follows:

$$T = t_1 + t_2$$

$$T = 50 + 150 = 200 \text{ microseconds}$$

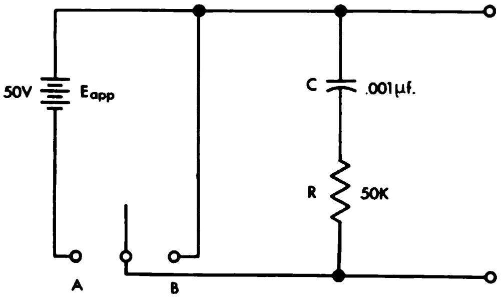
The frequency is found as before.

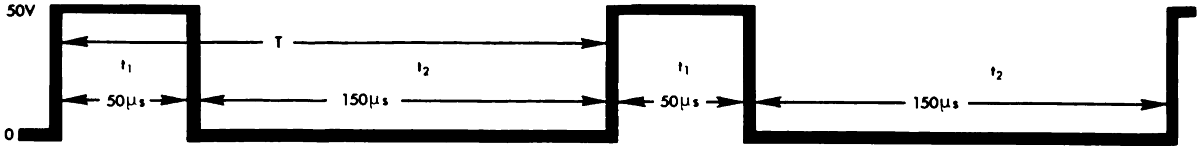
$$f = \frac{1}{T}$$

$$f = \frac{1}{200 \times 10^{-6}} = \frac{10^6}{200}$$

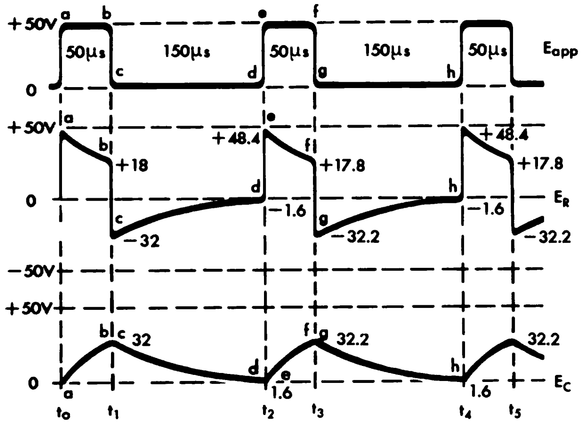
$$f = 5000 \text{ cps}$$

The universal TC chart is used in computing waveshapes E_R and E_C as before. However, t_1 is used when the capacitor is charging, and t_2 is used when the capacitor is discharging. RC is the same in both cases. The waveshapes for this circuit are shown in the illustration at the top of the next page.





Applying Rectangular Pulses to RC Circuit



Output Waveshapes When Rectangular Pulses are Applied

At the instant t_0, the switch is placed in position A. As shown in the diagrams E_R is 50 volts and E_C is zero. While +50 volts is applied for 50 microseconds the capacitor charges and the resistor voltage decreases. The voltages shown at point b are found as follows:

$$R \times C = (50000)\ (.001)$$

$$RC = 50 \text{ microseconds}$$

$$\frac{t_1}{RC} = \frac{50}{50} = 1$$

The capacitor charges to 63% of the applied voltage — 31.5 volts — in one TC. This is shown as 32 volts on the diagram. By using curve B or by subtracting 32 volts from E_{app}, you find E_R equals 18 volts.

At time t_1, the switch is thrown to position B. The capacitor becomes the source of voltage and an equal voltage with opposite polarity instantaneously appears across the resistor. As shown at point c, E_R is -32 volts and E_C is 32 volts. The capacitor discharges through the resistor for 150 microseconds. In this case, $\frac{t_2}{RC} = \frac{150}{50} = 3$. From curve B of the universal TC chart you find that E_R and E_C drop to 5% of 32 volts or 1.6 volts as shown at point d.

At time t_3, the switch is again thrown to position A. Instantaneously, the voltage on the resistor becomes $E_{app} - E_C$ which is equal to 48.4 volts as shown at point e. You may compute the values of E_C and E_R at times t_4, and t_5 in a similar manner.

DIFFERENTIATION AND INTEGRATION

The term *differentiation*, when applied to an electronic circuit, denotes the distortion of an input wave so that it yields an output as follows:

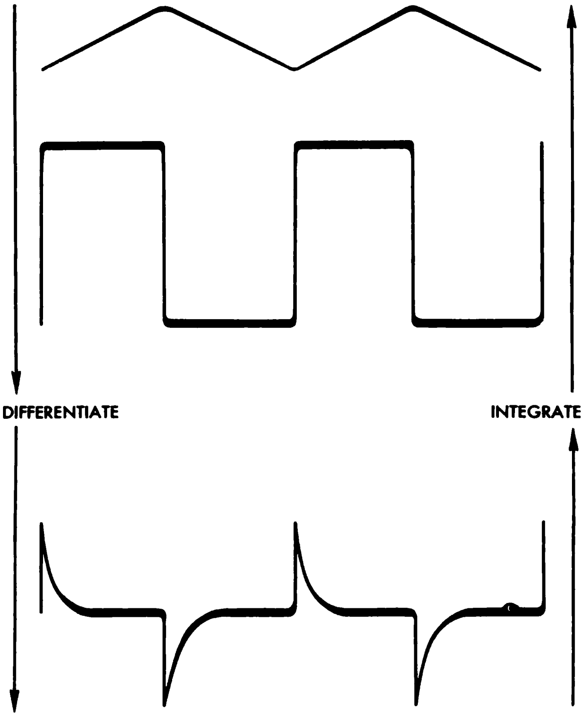
A triangular wave is distorted into a square wave; a square wave is distorted into a peaked wave.

Differentiation and *integration*, like addition and subtraction or multiplication and division, are reverse processes. Integration, when applied to an electronic circuit, denotes the distortion of an input wave so that it reverses the process of differentiation as follows:

A peaked wave is distorted into a square wave; a square wave is distorted into a triangular wave.

The illustration below, Differentiation and Integration, summarizes these results. To differentiate a waveshape, take the waveshape directly below it as the arrows indicate. To integrate a waveshape, take the waveshape directly above it as the arrows indicate.

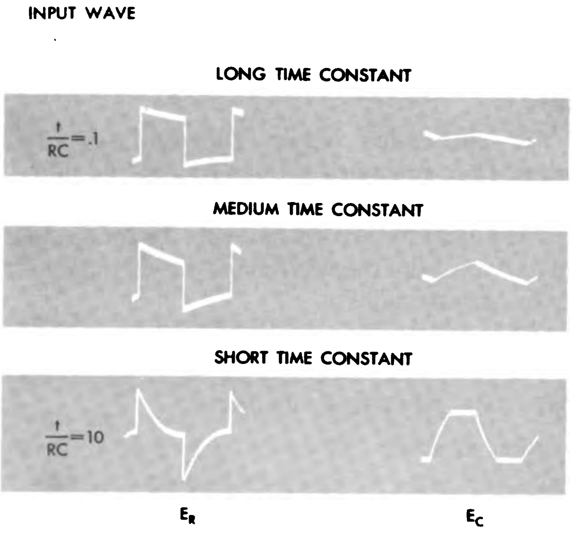
Differentiating or integrating a sine wave results only in a change in phase — the output waveshape remains a sine wave.



Differentiation and Integration

Long and Short Time Constants

Square and rectangular voltage waveshapes have been applied to RC circuits and the results have been observed. You noted that the values of R, C, and t affected the shape of the output waves E_R and E_C in each case. In the illustration below are shown output waveshapes obtained for three different values of t, R, and C. These waveshapes are for $\frac{t}{RC} = .1$, $\frac{t}{RC} = 1$, and $\frac{t}{RC} = 10$.



Output Wave Shapes When Square Wave Is Applied to RC Circuits with Different Time Constants

The variation in the waveshapes shown represent the results of the relationship between RC and t. If the time constant (RC) of a series RC circuit is ten (or more) times as long as t, it is called a *long time constant.* This relationship may be inverted as follows: If the time t is one-tenth (or less) of the product RC, the time constant is long. Expressed mathematically:

$\frac{t}{RC} = .1$ or less, denotes *a long time constant.*

A time constant which is equal to (or less than) one-tenth of t is called a *short time constant.* This relationship may be inverted as follows: If the time t is ten (or more) times as long as the product RC, the time constant is short. Expressed mathematically:

$\frac{t}{RC} = 10$ or more, denotes *a short time constant.*

The values for determining whether a time constant is long or short have been chosen arbitrarily. A medium time constant is one whose value is between the long and short time constant.

Note that the product of R and C alone does not determine whether a time constant is long or short. This is determined by the relationship that exists between the time t and the time constant. You may find that 10,000 microseconds denotes a short time constant for one input, while 50 microseconds denotes a long time constant for another.

Study the output waveshapes shown for each case and observe the following:

When the time constant is long $\left(\frac{t}{RC} = .1\right.$ or less), E_R has almost the same waveshape as the input with approximately the same amplitude; E_C is greatly distorted so that its shape is almost triangular and its amplitude is much less than that of the input.

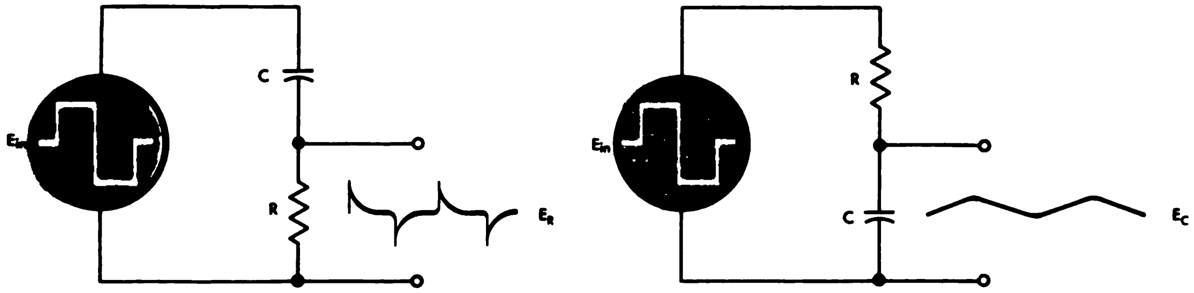
When the time constant is medium $\left(\frac{t}{RC} = \right.$ between .1 and 10), E_R is distorted and its peak-to-peak amplitude is greater than that of the input; E_C is distorted but not as much as for the long time constant, while its amplitude is slightly less than that of the input.

When the time constant is short $\left(\frac{t}{RC} = 10\right.$ or more), E_R is greatly distorted into a peaked waveshape and its peak-to-peak amplitude is twice the amplitude of the input square wave; E_C has almost the same waveshape as the input.

Differentiating and Integrating with RC Circuits

A circuit which differentiates an input is called a *differentiating* circuit. Since a differentiating circuit changes a square wave input to a peaked wave, it is also called a *peaking* circuit. An *integrating* circuit is one which integrates an input wave. It therefore changes a square wave input to a triangular wave.

Although they were not called by these names, differentiating and integrating circuits have already been studied. Examine the illustration on this page. Note that the output waveshape E_R



DIFFERENTIATION
WITH SHORT TIME CONSTANT

INTEGRATION
WITH LONG TIME CONSTANT

Differentiating and Integrating RC Circuits

for a short time constant is a peaked wave and that the output waveshape E_C for a long time constant is a triangular wave. These are reproduced in the illustration shown above, Differentiating and Integrating RC Circuits. A square wave is shown as the input in each case; however, these circuits will differentiate or integrate any type of input. If the input to the differentiating circuit is changed to a triangular wave, the output is a square wave; if the input to the integrating circuit is a peaked wave, the output is a square wave. A sine wave applied to either circuit gives an output which is a sine wave whose phase has been shifted.

In many cases, a circuit is needed which will not distort the input. Examine again the illustration on page 296. Note that, when the time constant is long, the output E_R is almost the same as the input. Such a circuit, in which the output is taken across the resistor, is called a coupling circuit. RC coupling circuits are common in radio and radar applications. An RC coupling circuit and an RC integrating circuit are identical; however, the component across which the output is taken differs.

To summarize:

1. An RC circuit is a differentiating circuit if the time constant is short and the output is taken across the resistor.

2. An RC circuit is an integrating or coupling circuit if the time constant is long. If the output is taken across the capacitor, it is an integrating circuit; if the output is taken across the resistor, it is a coupling circuit.

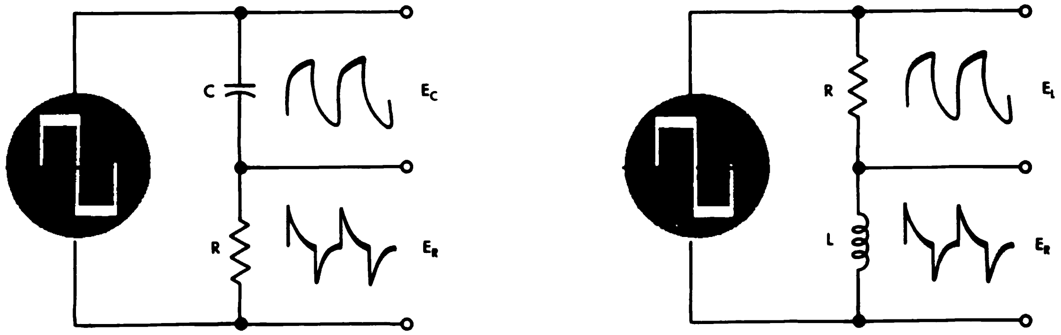
RL CIRCUITS

The effect of transients on RL circuits is similar in many ways to the effect of transients in RC circuits. In your study of the RC circuit, you observed that a capacitor offered opposition to a sudden change of *voltage* across it. The voltage across the capacitor built up from zero to maximum in approximately 5 time constants. In an RL circuit, on the other hand, the inductance offers opposition to a change in *current*. The current in an RL circuit therefore builds up from zero until it reaches a maximum. The rate at which current changes through a coil depends upon the voltage across it and the inductance of the coil. As current builds up in an RL circuit, the drop in voltage across the resistor causes a decrease in the voltage applied to the coil and, therefore, a decrease in the *rate* at which the current builds up. This decrease in the *rate* of increase in current continues until the current becomes constant. When the current reaches a steady value, it has maximum value $\left(\frac{E_{app}}{R}\right)$. Note that the description of the rise of current in an RL circuit is similar to the rise of voltage across the capacitor in an RC circuit. Curve A of the universal TC chart may therefore be used to show this rise of current.

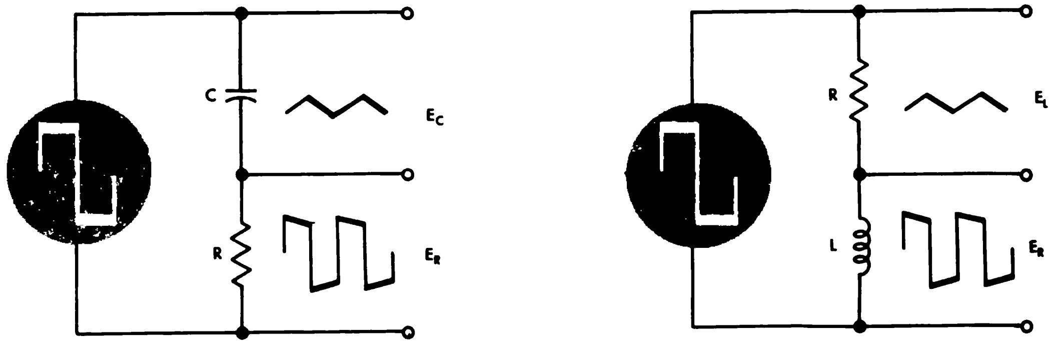
Time Constant of an RL Circuit

The shape of the curve that shows the buildup of current in an RL circuit has been described. What factors affect the time necessary for this buildup? In the RC circuit, you observed that the voltage E_C reached a maximum in approximately 5 RC. In the series RL circuit, the components (L and R) also have an effect. The rate at which current changes through a coil depends upon the value of the inductance of the coil. The greater the inductance, the less is the rate of change of current and the greater is the time needed for the current to build up to a maximum. The time required for the current to build up is therefore directly proportional to the magnitude of the inductance.

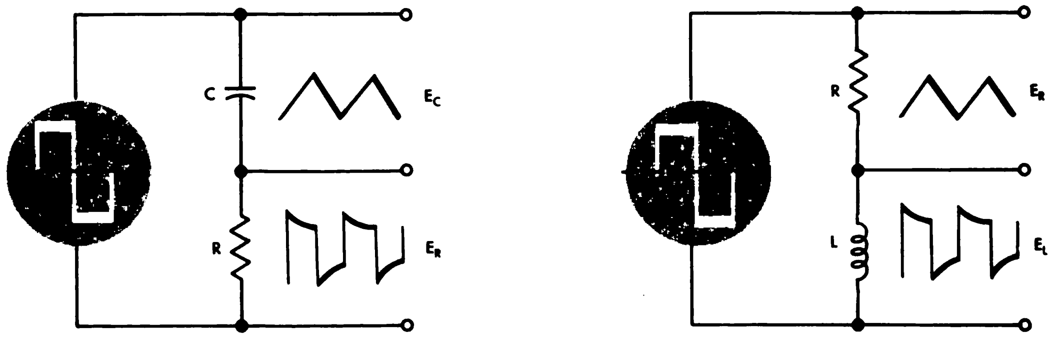
The maximum value $\left(\frac{E_{app}}{R}\right)$ of current in an RL circuit is determined by the magnitude of the resistance present. The greater the value of the resistance, the smaller is the value of the maximum current and the time needed for this



SHORT TIME
CONSTANT CIRCUITS



LONG TIME
CONSTANT CIRCUITS



MEDIUM TIME
CONSTANT CIRCUITS

Square Wave Applied to RC and RL Circuits maximum to be reached. The time required for the current to build up is, therefore, inversely proportional to the magnitude of the resistance.

Since the time needed for current in an RL circuit to reach a maximum is directly proportional to the inductance and inversely proportional to the resistance, the time is proportional to the product of L and $\frac{1}{R}$. This product $\left(\frac{L}{R}\right)$ represents a period of time similar to the product of R and C in an RC circuit and is called a time constant. Expressed mathematically:

$$TC = \frac{L}{R}$$

where TC is the time constant in seconds, L is the inductance in henries, and R is the resistance in ohms.

At the end of one TC in an RL circuit, the current has reached 63% of its maximum value. At the end of two TC, the current has reached 86.5% of its maximum value. At the end of five TC the current is approximately maximum. The current buildup in the RL circuit may therefore be represented by curve A of the universal TC chart if the quantity $\frac{L}{R}$ is used as a time constant.

The universal TC chart may, therefore, be used with RL circuits in the same manner as with RC circuits. In this case, the number of TC along the horizontal axis is found as follows:

$$\text{No. of TC} = \frac{t}{TC} = \frac{t}{L/R}$$

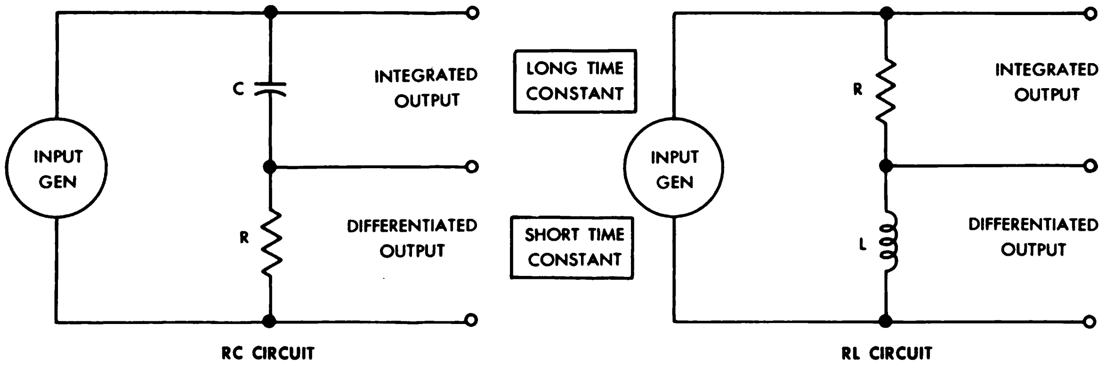
$$\text{No. of TC} = \frac{Rt}{L}$$

The percentage along the vertical axis represents the percentage of maximum current flowing at each instant.

Applying a Square Wave

When a square wave or a rectangular pulse of voltage is applied to an RL circuit, the output waveshapes are similar to those obtained with RC circuits. Some typical output waveshapes for an RL circuit are shown on page 298. The values at different points of the waveshape may be computed — as for RC circuits — by using the universal TC chart.

The value of $\frac{Rt}{L}$ determines whether the time constant is long, medium, or short. For $\frac{Rt}{L}$ $=0.1$ or less, the time constant is long. For $\frac{Rt}{L}$ $=1$, the time constant is medium. For $\frac{Rt}{L}=10$ or more, the time constant is short. Differentiation and integration are accomplished by RL circuits; however, the differentiated output is taken across the inductance when the time constant is short, and the integrated output is taken across the resistor when the time constant is long.



Differentiated and Integrated Outputs

SUMMARY

The most important characteristics of RC and RL circuits may be summarized as follows:

1. RC and RL series circuits are used in radar circuits as waveshaping or wave-controlling devices.

2. Ohm's and Kirchhoff's laws apply to RC and RL circuits.

3. A time constant for an RC or RL circuit represents the time required for the circuit to complete approximately 63% of its transient action. It is equal to R×C in an RC circuit and $\frac{L}{R}$ in an RL circuit.

4. The Universal Time Constant Chart may be used to determine the per cent of transient action completed in any specified number of time constants for RC and RL circuits.

5. The number of time constants in a specified time is $\frac{t}{RC}$ in an RC circuit and $\frac{Rt}{L}$ in an RL circuit.

6. All the transient actions in RC and RL circuits (i. e. condenser charging, current rise, etc.) follow exponential curves.

7. A circuit has long time constant if $\frac{t}{RC}$ or $\frac{Rt}{L}$ is equal to or less than 0.1.

8. A circuit has short time constant if $\frac{t}{RC}$ or $\frac{Rt}{L}$ is equal to or more than 10.

9. The diagrams on page 299 indicate where the integrated and differentiated outputs are taken in RC and RL circuits.

10. The effect of differentiation and integration on waveshapes is shown on page 295. To differentiate a waveshape, take the waveshape directly below it; to integrate a waveshape, take the one above.

Chapter 31

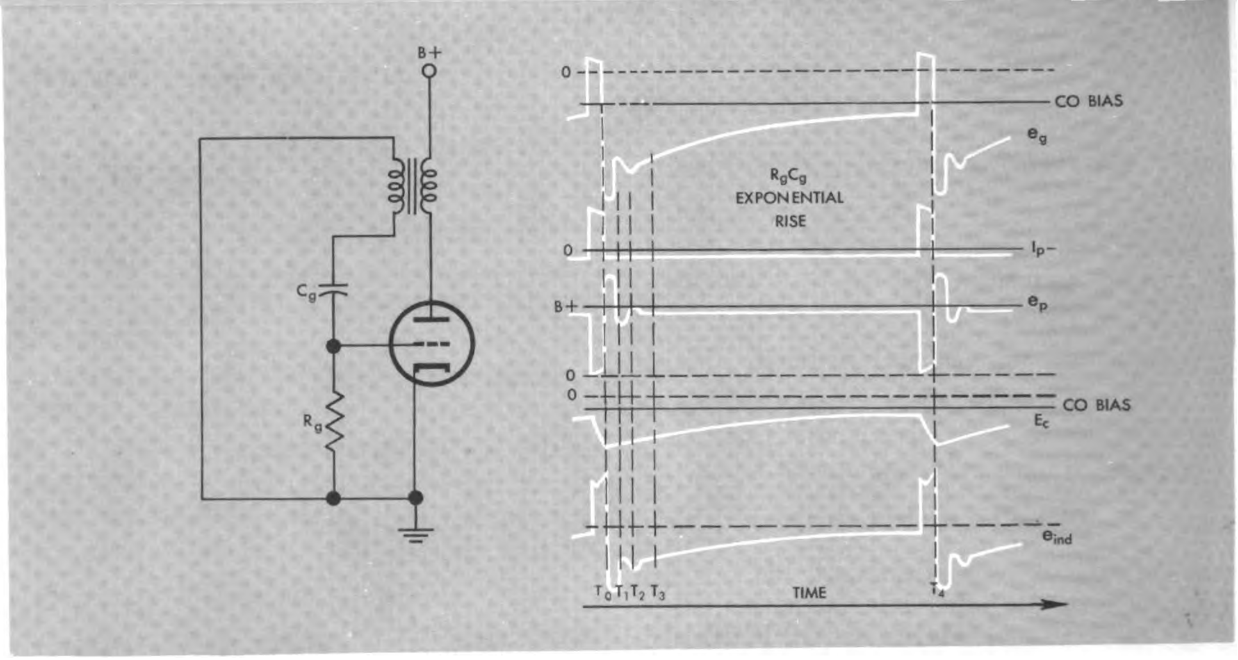
Blocking Oscillators

The use of narrow pulses in various radar circuits is very common. Pulses that have been synchronized and whose frequency is controlled are used in radar to accomplish the synchronized triggering of many circuits. To effect the accurate measurement of the distance to a target, it is essential that the sweep circuit and the transmitted pulse be synchronized, that is, triggered at the same time or definitely related times.

Narrow pulses may be used to synchronize the sweep circuit with the transmitter pulse. The narrow pulses are applied to different circuits causing them to operate simultaneously. Regardless of how the pulses are applied or regardless of their polarity, a circuit which generates very narrow pulses at a controlled frequency is a necessity in modern radar equipment. One special circuit that produces very narrow pulses is the *blocking oscillator*.

THE FREE RUNNING BLOCKING OSCILLATOR

A blocking oscillator may be either free-running or synchronized. A free running blocking oscillator generates a series of narrow pulses at a frequency determined by the constants of the circuit. A synchronized blocking oscillator generates a series of narrow pulses at a frequency determined by the frequency of the synchronizing signal. The frequency of the output pulses in such cases is either the same as the synchronizing signal or a submultiple of it.



A Basic Blocking Oscillator and Waveshapes

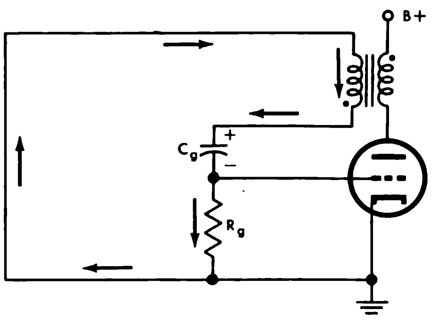
The circuit for the conventional blocking oscillator is shown in the illustration above. It consists of a triode having the primary of a step down transformer (approximately 3:1) in its plate circuit. The secondary is coupled through a capacitor to the grid. For the purpose of understanding its operation, assume that the grid capacitor C_g has been negatively charged by a preceding cycle, as shown in the illustration below.

The tube, therefore is biased well below cut off. As the charge on the capacitor leaks off, the biasing voltage is reduced to the point where the tube begins to conduct. As plate current starts to flow, a magnetic field is set up around the plate winding of the transformer. The dots at each winding indicate similar polarities. If, for example, a current flows through a winding so that the dot end is positive, the field set up in the core induces voltages in the other winding that makes the dot end in this winding positive at the same time. As the plate current flows in the primary it causes a voltage to be induced in the secondary with the polarity as shown in the illustration below.

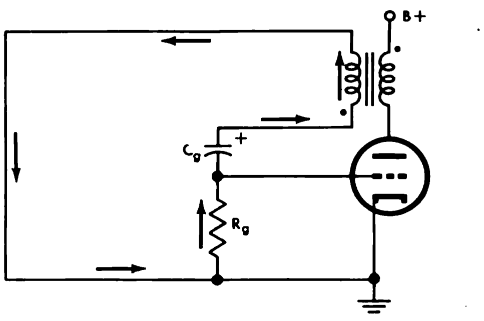
The polarity of the induced voltage is such as to make the grid more positive, causing a greater current through the primary. The induced voltage caused by this change makes the grid still more positive, the plate current still greater, and drives the plate more negative. This action is cumulative and occurs almost instantaneously, as the waveshapes in the illustration above indicate.

Analysis of Waveshapes

The tube begins to conduct at time t_o. Note that the plate current reaches a maximum very quickly, the grid becomes positive, and the plate is driven in a negative direction from B+ toward



Discharge Path of the Capacitor



Charge Path of the Capacitor

zero. As soon as plate current reaches an unchanging maximum value, you may expect the induced voltage to become zero. This is exactly what would happen if no current flowed in the secondary. The induced voltage, however, is of such polarity as to make the grid draw current. This grid current flows through the secondary of the transformer. The magnetic flux of the transformer core therefore does not reach its equilibrium value as soon as the primary current becomes a maximum. In fact, the induced voltage wave in the illustration indicates there is a rise after a small initial drop between times t_o and t_1. This rise is caused by the counter EMF induced in the secondary as current flows from the grid. This secondary current, however, charges the capacitor, C_g; hence, this current is continuously decreasing until it finally becomes zero when the capacitor is fully charged. This decrease in current does not vary at an exponential rate like that which occurs in the R-C circuits already studied, because both the applied voltage (the induced voltage in this case) and the grid to cathode resistance of the tube are varying. In fact the current decreases at an almost linear rate. The capacitor charges at a rate of charge that is almost linear as shown by the waveshape E_C between times t_o and t_1.

Note that in the illustration showing the charge path of the capacitor the voltage e_g is the algebraic sum of the induced voltage and the voltage E_C. Hence, the voltage e_g may be obtained by adding the induced voltage waveshape and E_C. The resultant waveshape e_g indicates a drop in grid voltage from t_o to t_1. The plate current follows this grid voltage and drops from its maximum value while e_p rises. This continues until t_1 is reached. At that time the capacitor is charged and the secondary current is zero. Grid voltage (e_g) and plate current (i_p) become steady and the voltage induced in the secondary becomes zero. Instantaneously, the voltage on the grid is driven in a negative direction as the capacitor becomes the only source of voltage across the grid resistor R_g. The sudden decrease in plate current flowing through the tube caused by the drop in voltage on the grid causes a current change through the primary in the opposite direction. Hence, the induced voltage also reverses polarity. This induced voltage aids the voltage across the condenser and acts as an additional source of voltage across the resistance R_g. Instantaneously the voltage across R_g drops much below the cutoff value; the magnitude of this instantaneous drop is equal to the sum of the voltage across the capacitor and voltage induced in the secondary. This effect is shown in the waveshapes at time t_1. At this time, grid voltage drops much below cut off, the plate current becomes zero, and the plate voltage rises to B+. When the grid voltage has reached a maximum negative value, the capacitor C_g begins to discharge through R_g at an exponential rate. The grid voltage e_g begins to rise along an exponential curve whose rate of rise depends upon the values of R_g and C_g. This exponential rise continues until the voltage across R_g reaches the cutoff bias at time t_4. From this time, the action is repeated with a consequent repetition of waveshapes across the various components as shown in the illustration.

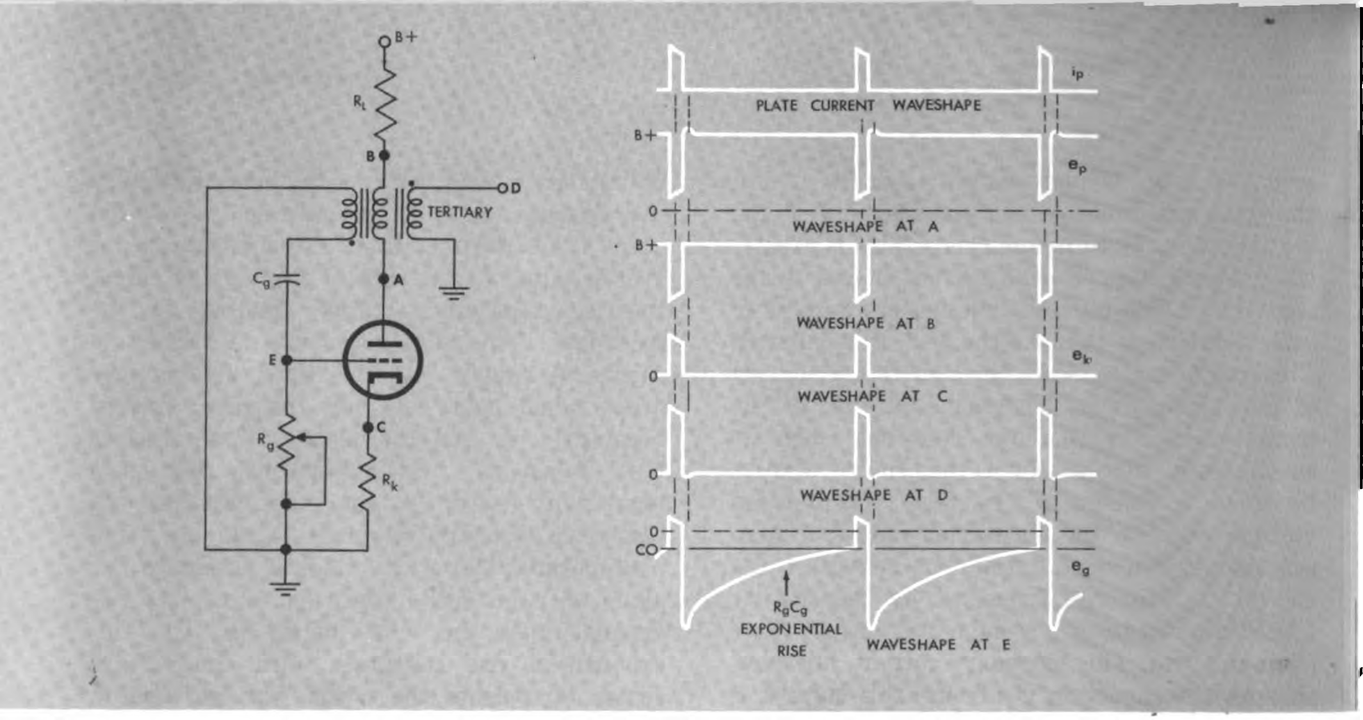
You will also notice in the illustration that e_g does not follow an exponential curve between times t_2 and t_3. This is due to an oscillation set up in the primary winding by its inductance and distributed capacitance, when the tube is suddenly cutoff. This oscillation is reflected into the grid circuit by the secondary.

Factors That Affect the Frequency of the Blocking Oscillator

If an output is taken from plate to cathode, the waveshape e_p represents the output wave of the blocking oscillator. If you disregard the damped oscillations which occur, the output may be considered as a series of negative-going pulses generated each time the tube conducts, followed by a period of nonconduction, with the plate voltage at B+. Since the time during which the tube conducts is of very short duration (approximately 1 μs) the time of nonconduction may be considered as the time (period) of a complete cycle. Hence, the frequency, which is the reciprocal of the period T, is dependent upon the period of nonconduction. The factors that determine how long the tube is nonconducting (cutoff) are:

1. The critical drop which occurs immediately after conduction and which is reflected as a negative voltage on the grid.

2. The magnitude of the damped oscillations which occur immediately after the initial drop.



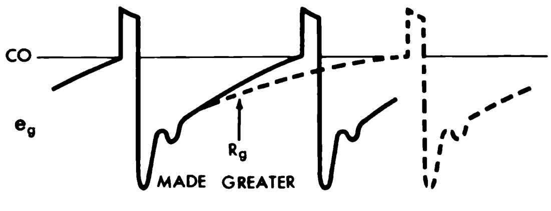
Blocking Oscillator with R_g Variable and Waveshapes

3. The cutoff value of the tube.

4. The values of R_g and C_g, which determine the exponential rise of the grid waveshape toward cutoff value.

The first three factors depend upon the value of B+, the characteristics of the tube, and the transformer used. Once these circuit components are chosen, they are not readily variable as a means of controlling the frequency of the pulses. The fourth factor depends upon the values of C_g and R_g which may readily be made variable. R_g is usually made variable to control the frequency of a blocking oscillator, as shown in the illustration above.

The effect of increasing R_g is shown in the illustration below. As R_g is increased the charge on C_g leaks off more slowly and the tube remains cut off for a longer period. C_g is not usually changed to vary the frequency of the blocking oscillator as it has an effect on the width of the pulse. The smaller the capacitor the less time it takes to charge, the less time it takes for the plate current to reach a steady state and for the polarity to be reversed, therefore, the narrower the pulse.



Effect of Varying R_g on Frequency

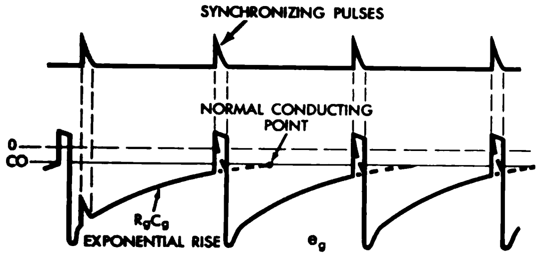
Notice in the illustration above, that in addition to making R_g variable, some components have been added to the blocking oscillator circuit.

1. R_k, a cathode resistor whose value is low and across which an output can be taken.

2. R_L has been placed between B+ and the primary winding to help suppress the undesirable oscillations. In some cases, resistances of small value are shunted across the transformer windings to suppress oscillations.

3. The transformer has three windings: primary, secondary, and a tertiary, across which an output may be taken.

The output waveshapes of this blocking oscillator with undesirable oscillations suppressed are also shown in the illustration above. Notice that you have a choice of polarity of output waveshape: positive from points C, D, E and negative at points A and B. Also notice that the resistor R_1 has reduced the oscillations occurring in the plate with a resulting smoother voltage rise on the grid.



Synchronizing a Blocking Oscillator

Synchronizing a Blocking Oscillator

The natural frequency of the blocking oscillator varies somewhat from time to time due to aging components, effects of temperature, etc. Even after a thorough warm up, it exhibits a tendency toward instability. Synchronizing pulses are often used to stabilize the frequency of the blocking oscillator. This is generally accomplished by applying positive synchronizing pulses to the grid of a blocking oscillator to cause it to conduct at a time slightly earlier than normal. In the illustration above positive pulses are being applied to the grid of the blocking oscillator.

Notice that the grid voltage e_g is the sum of E_C plus the positive synchronizing pulse. Without the synchronizing pulse, the tube would not conduct until the grid voltage reached the point of cutoff. However, with the synchronized pulse applied to the grid the tube conducts earlier, since at this time the sum of e_g plus the synchronizing pulse is sufficient to drive the grid above cutoff and cause the tube to conduct. The frequency of blocking oscillator is now the same as that of the synchronizing pulse.

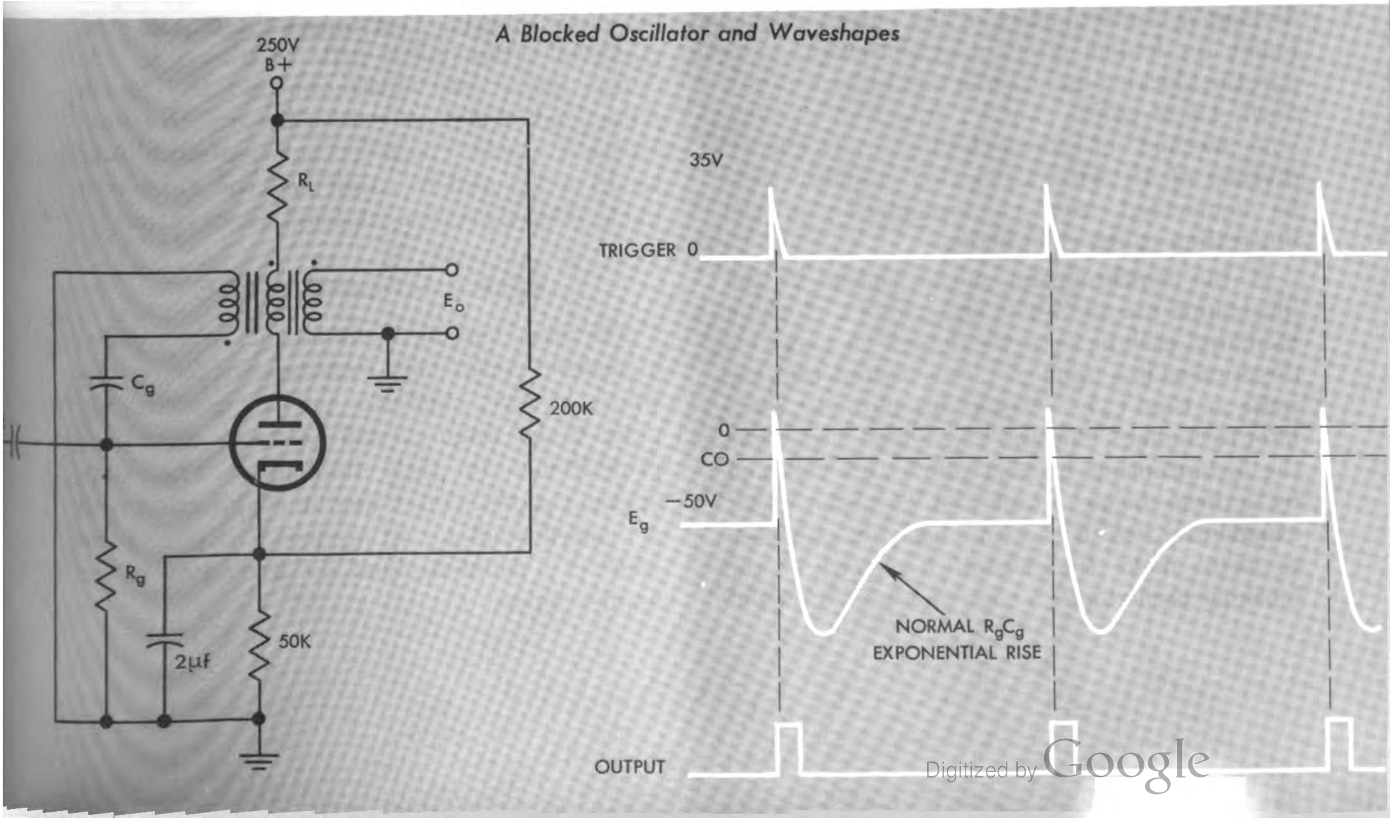
Blocking oscillators may also by synchronized by applying negative pulses to the cathode or by coupling the synchronizing pulse through a tertiary winding of the plate transformer to the grid. You will encounter all these methods for synchronizing a blocking oscillator.

THE BLOCKED OSCILLATOR

Some radar circuits require a blocking oscillator which operates only when triggered. Such a circuit is called a blocked oscillator since a permanent DC voltage keeps the tube cut off and prevents the circuit from oscillating on its own accord. A blocked oscillator is shown in the illustration below.

Cutoff is maintained by a voltage divider from the plate supply voltage. With the values shown, the cathode voltage will be 50 volts positive with respect to ground. The grid is grounded and is therefore 50 volts negative with respect to the cathode. Cutoff for many of the tubes used in this type circuit occurs at about −18 volts, so the trigger pulse must be more than 32 (50−18) volts in amplitude to start a cycle of operation. In the grid waveshapes, the normal discharge of

A Blocked Oscillator and Waveshapes



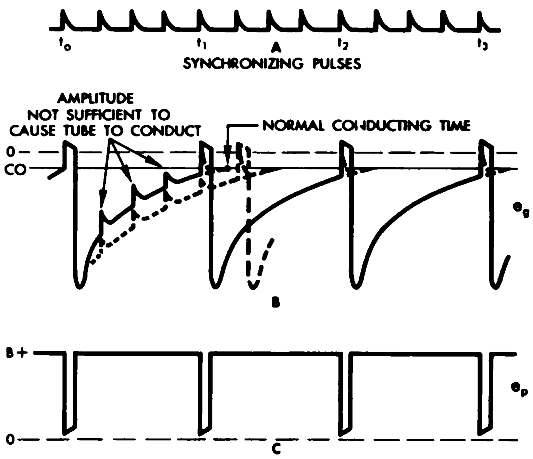
the grid condenser reduces the grid voltage after a cycle of operation, but the discharge is to a −50 volts, so the circuit cannot again conduct without another trigger pulse. The important waveshapes of a blocked oscillator also appear in the illustration on page 305.

The positive pulse must be applied to the grid only. It is not practical to apply a trigger pulse to the cathode because of the filter condenser that is normally found there. A negative pulse may be applied to the plate where the transformer will invert the pulse to a positive one on the grid, or a positive pulse may be applied through a tertiary winding of the plate transformer where it will be applied as a positive pulse on the grid. The output of this circuit may be negative, taken directly from the plate. It may also be taken from a tertiary winding where a pulse similar to the plate pulse can be obtained with either polarity, depending on which side of the tertiary is grounded.

A BLOCKING OSCILLATOR AS A FREQUENCY DIVIDER

Although the frequency stability of a blocking oscillator is improved when it is synchronized with incoming pulses, the frequency of the input and that of the output need not always be equal. In many cases, the output frequency may be a submultiple of the incoming frequency. In each of these cases, the blocking oscillator is said to be counting down or acting as a frequency divider. If the frequency of the output is one-fourth the frequency of the synchronizing voltage, the blocking oscillator is said to be counting down 4:1 or acting as a 4:1 frequency divider. The method of accomplishing a frequency division of 4:1 with a blocking oscillator may be understood from an examination of the illustration to the right above.

If the synchronizing pulses shown in the illustration are applied to the grid of a blocking oscillator, each pulse drives the grid in a positive direction as it is applied. Let us assume that the pulse at time t_0 has caused the blocking oscillator to conduct. The grid waveshape then shows the instantaneous increase above zero followed by a drop below cutoff and then the usual exponential rise. If the time constant is chosen so as to make the natural frequency of the blocking oscillator less than the desired output (one-fourth the input frequency) and, at the same time, the amplitude of the synchronizing voltage chosen so as to be sufficient to drive the tube above cutoff only on each fourth pulse, the desired result is achieved. The grid is driven above cutoff on each fourth pulse, and output pulses occur at a frequency which is one-fourth that of the input frequencies. Thus the output pulses occur at t_0, t_1, t_2, and t_3.



Waveshapes of a Blocking Oscillator Acting as a Frequency Divider

By making R_g variable, the above blocking oscillator may be made to count down exactly as desired. In order to decrease the output frequency (count down ratio of 5:1), R_g is increased slightly. This causes the exponential curve to flatten out a little as shown by the dotted line in the illustration and the fourth pulse is not sufficient to drive the grid to cutoff. However, the fifth pulse is of sufficient amplitude to drive the grid to cutoff and the tube conducts. This gives us a 5:1 count-down ratio.

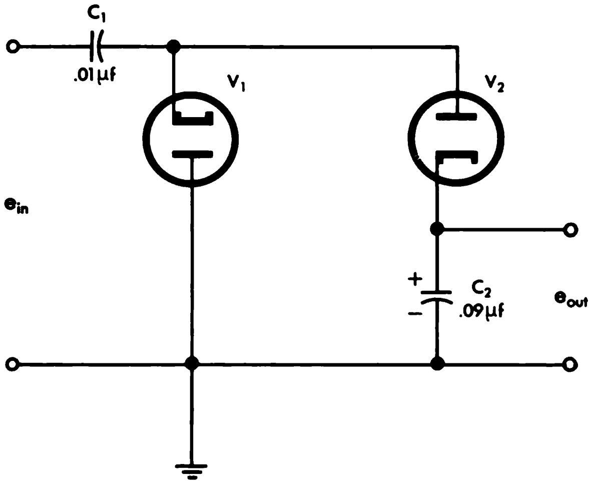
STEP-CHARGING DIODE

The step-charging diode circuit, often called a *counting circuit* or a *frequency divider*, is one which receives uniform pulses, representing units to be counted, and produces a voltage proportional to their frequency. The step-charging diode is often used in conjunction with a blocked oscillator to produce a trigger pulse which is a submultiple of the frequency of the applied pulses.

Circuit Analysis

In the step-charging diode circuit shown in the illustration to the right, 100 volt peak-to-peak square waves are being applied. C_2 has a capacitance of .09 microfarad and C_1, .01 microfarad. Therefore, since the condensers are in series, C_2 will charge to 10% of each input pulse and C_1 to 90% of each input pulse (the voltage distribution in condensers in series is in inverse ratio to the capacity of the respective condensers).

The charge on condenser C_2 with the first 100 volt pulse applied will be 10 volts, while C_1 will charge to 90 volts. When the first pulse ends, C_1 discharges through V_1, but C_2 holds its 10 volt charge because it has no discharge path. The second input pulse now occurs. The 10 volt charge on C_2 opposes the 100 volts of the second pulse, so effectively only 90 volts is applied to the circuit. C_2 will again charge to 10% of the applied voltage (9 volts). This is in addition to the initial 10 volts, so at the end of the second pulse, the C_2 voltage is $10+9$, or 19 volts. C_1 will charge to the remaining 81 volts, but again, at the end of the pulse it will discharge through V_1. The third pulse will be 100 volts again, but 19 volts will oppose it. Therefore, only $100-19$, or 81 volts, is being applied effectively. C_2 again will charge to 10% of this value, or 8.1 volts. The voltage on C_2 is now $19+8.1$ or 27.1 volts.



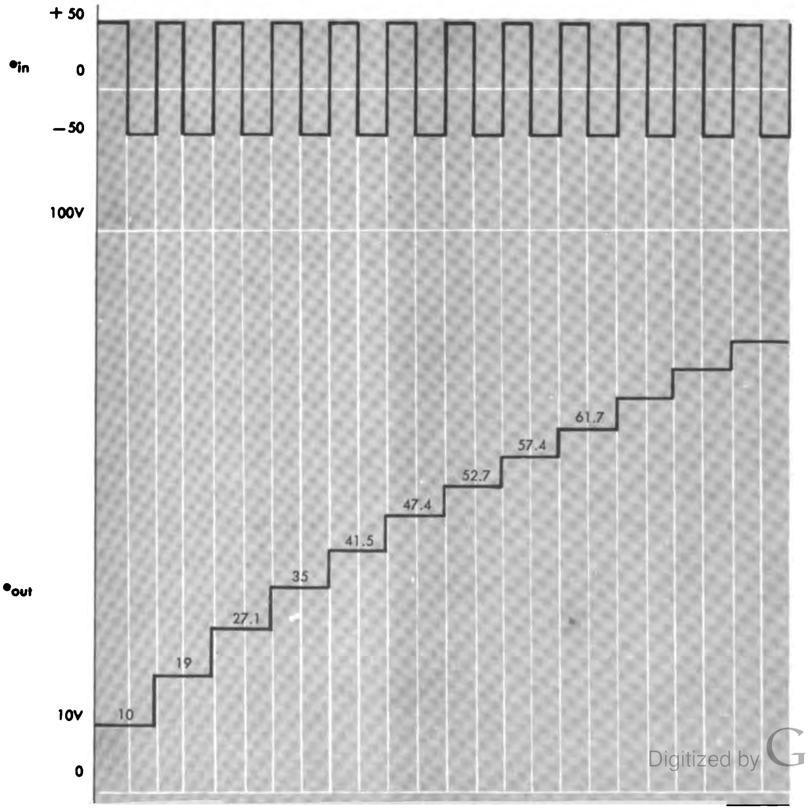
Step Counting Circuit

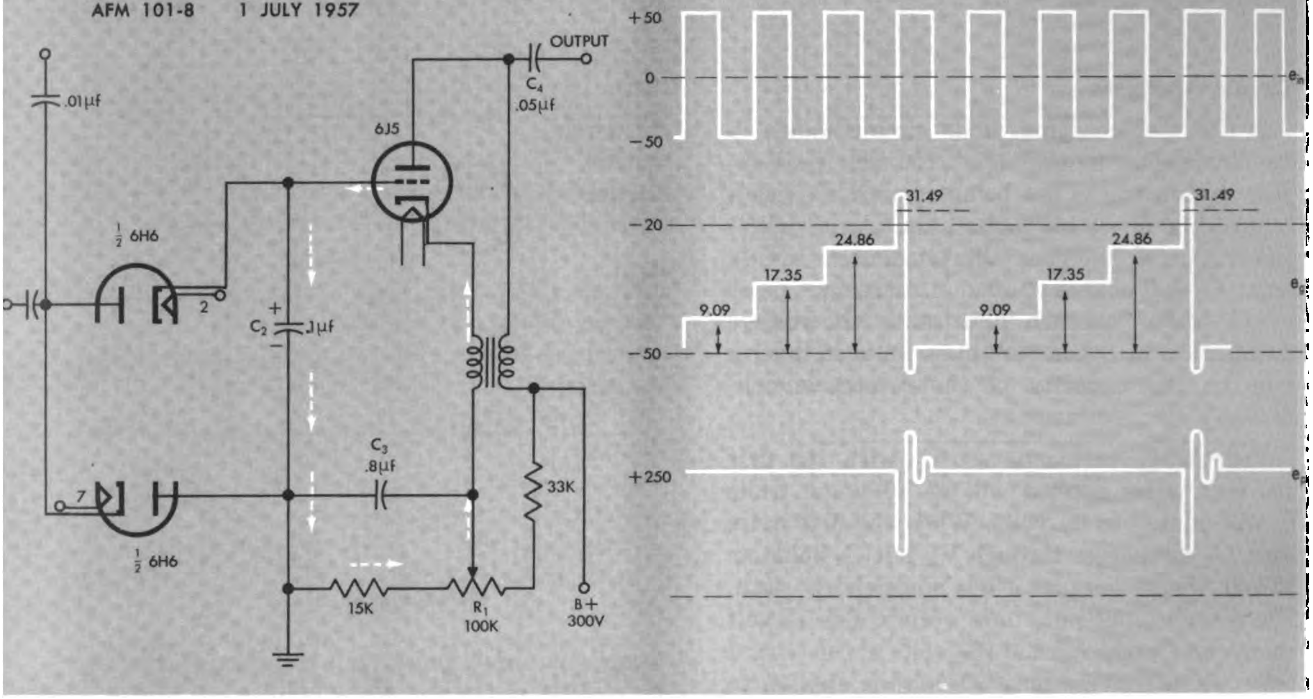
Successive input pulses will continue to charge C_2 to 10% of the remaining voltage until finally it becomes charged to 100 volts. The buildup of the charge on C_2 is shown in the illustration below.

Use of a Counting Circuit

The illustration on the next page shows a step-counting circuit followed by a blocked oscillator on which the fixed bias is 50 volts. The cut off voltage for this tube is −20 volts. This blocked oscillator functions in exactly the same way as the one shown previously. The only difference is

Voltage Buildup on Capacitor C_2





Step-Charging Diode and Waveshapes

that the feedback is in the cathode circuit rather than the grid circuit.

If the tube will not conduct until the grid reaches 20 volts negative with respect to the cathode, the grid potential must be increased by 30 volts before the blocked oscillator will conduct, and produce an output pulse.

Since C_2 charges to one-eleventh of each applied pulse, you can see from the illustration that four input pulses must be applied to the circuit before V_3 will conduct giving us an output pulse. The circuit is then acting as a frequency divider, by counting the input frequency down at a ratio of 4:1. When V_3 conducts, C_2 immediately discharges through the tube as shown by the dotted line, and another cycle of operation starts.

To cause the circuit to divide by some other ratio, the bias is set at a different value by moving the arm of R_1. For example a fixed bias of 32 volts would allow V_3 to conduct on the 2nd step, making the circuit a 2:1 frequency divider.

Chapter 32

Nonsinusoidal Generators and Waveshaping Circuits

The blocking oscillator, as you have seen, is a nonsinusoidal generator used to generate voltage pulses of short time duration for use as triggering pulses. Often it is desirable to obtain rectangular voltage pulses of relatively long time duration (several hundred or thousand microseconds). A pulse of such long duration is useful to gate circuits; that is, allow a circuit to conduct or be cut off for the duration of the pulse. A nonsinusoidal generator which will produce pulses of long duration of positive or negative polarity is called a square-wave generator.

SQUARE-WAVE GENERATORS

There are various types of square-wave generators which you will encounter in your study of radar circuits. In this chapter, three of the most common types of square-wave generators (or multivibrators) are presented. These are the *plate-coupled*, the *start-stop* and the *Eccles-Jordan* multivibrators.

Plate Coupled Multivibrators

The plate coupled multivibrator is the simplest of the multivibrators. It produces almost a square wave when properly designed, and a perfect square wave when additions are made to the basic circuit. In the schematic on the next page, note that basically the multivibrator consists of two triode amplifier stages, with the output of the second stage introduced back into the input of the first. The phase inversion through the circuit is 360 degrees and any slight change in the first tube will be amplified through the circuit and cause a greater change in the same direction in the first tube. That is, feedback is *in phase* and the circuit oscillates.

When voltage is applied to the circuit, current flows in both tubes. If the circuit components are the same in both tube circuits and both tubes are of the same type, the circuit is said to be *symmetrical*, and both currents should be equal. But it is impossible to obtain a perfect balance in the

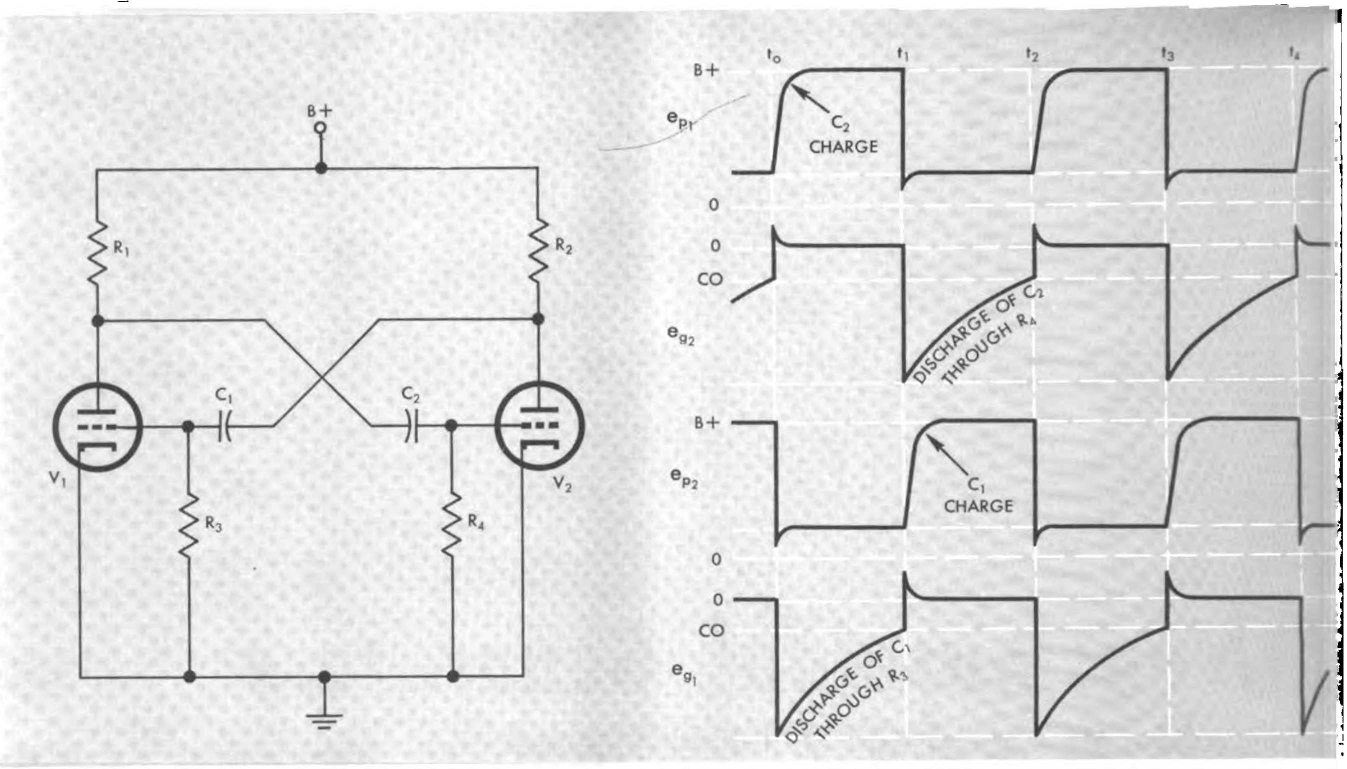


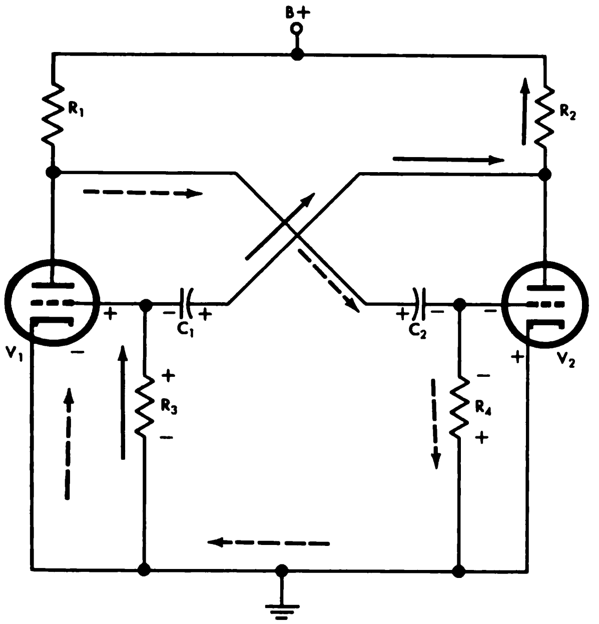
Plate-coupled Multivibrator and Waveshapes

circuit, because of the nonuniformity of available parts and the irregular way in which emission occurs from the cathode. Any slight change in current in one tube will start oscillations producing the output waveshapes shown.

WHEN V_1 CONDUCTS. Assuming that a slight current increase has occurred in the tube V_1 of the multivibrator, let us analyze the action with the aid of the diagram shown. The following series of events occur in rapid order: The increase of current in V_1 increases the voltage drop across R_1 which lowers the plate-to-ground voltage at V_1. This lower voltage is now less than the charge on C_2, and the condenser C_2 discharges toward the new plate voltage as shown by the dotted lines. The discharge current flows through R_4 making the grid end negative. The negative grid of V_2 decreases the current through V_2. This reduces the plate current through V_2 and the drop across R_2, causing the plate-to-ground voltage at V_2 to be higher. Condenser C_1, between this plate and ground, therefore charges to the new higher voltage. The charging current flows through R_3 toward the plate supply voltage as shown by solid arrows in the illustration. This current causes the top of R_3 and, therefore, the grid of V_1 to become more positive. The positive grid causes the original slight increase in current to become a tremendous current increase. As the changes go around the circuit, a great voltage drop occurs at the plate of V_1, the voltage on the grid of V_2 is decreased more, the plate voltage of V_2 goes up, making the grid of V_1 more positive, and further increasing the V_1 plate current.

This is a cumulative process which continues until the current in V_2 is decreased to zero by a grid voltage far below cutoff. With zero current, the plate voltage of V_2 becomes equal to B+ plate supply voltage, since there is no voltage drop across R_2. C_1 quickly charges to this value because of the grid current flow in V_1. The current through R_3 is now zero; so the grid-to-cathode voltage is also zero. The plate voltage of V_1 remains at a constant low value while the grid-to-cathode voltage is zero; so the condenser C_2 continues to discharge to this value through R_4. While C_2 is discharging, the current (dotted lines) holds V_2 cutoff. Everything remains at a standstill in the circuit except for the slow discharge of C_2.

The slow discharge of C_2 continues and the voltage across C_2 becomes lower and lower. The voltage across R_4 which equals the voltage across C_2, in accordance with Kirchoff's voltage law, also decreases. Upon reaching a voltage just less than the cut off voltage for V_2, a slight current starts flowing through V_2. This ends the period of conduction of V_1 and nonconduction of V_2. Swift changes follow this slight current, as described in the following paragraphs.



Action in Circuits When V_1 Conducts

WHEN V_2 CONDUCTS. The current through V_2 lowers its plate voltage. C_1 starts to discharge through R_3 (shown dotted in the illustration below), making the grid of V_1 more negative. The decrease in plate current which results raises the plate voltage of V_1. C_2 now recharges to the new V_1 plate voltage through R_4. The charging current makes the grid of V_2 more positive and the current through V_2 is increased tremendously. This regenerative process continues with grid of V_1 going negative and with the grid of V_2 going positive until V_1 is cut off by the discharge of C_1. At this point, V_1 plate voltage rises quickly to the plate supply voltage through V_2 grid current. At the same time, the plate voltage of V_2 becomes steady at a low value. This ends the rapid changes during condition one.

A second inactive period occurs as V_1 is held nonconducting by the slow discharge of C_1. The discharge continues until the C_1 voltage is just below cutoff for V_1. The slight current in V_1 builds up in the same manner as described for V_2 until V_1 conducts heavily and V_2 is cut off.

SUMMARY. In the waveshapes in the basic circuit diagram on page 310, the time between t_1 and t_2 is the time during which V_2 is cut off. The exponential grid voltage of e_{g2} shows how long the condenser C_2 holds V_2 nonconducting. Time t_2 is the time during which the rapid change described occurs. The half cycle following t_2 is the time during which V_1 is cut off.

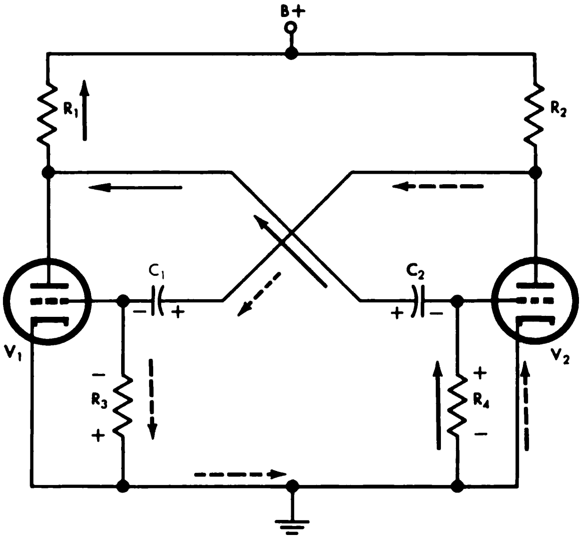
The operation of this multivibrator circuit consists of long periods of time when one tube conducts heavily while the other tube is cut off. This period is followed by an extremely rapid change to the other tube conducting heavily and the first tube cutoff. No external trigger pulses are required to make the circuit operate; therefore, it is termed a *free-running* multivibrator. Like the free-running blocking oscillator, however, the multivibrator may be synchronized at a higher frequency. Synchronizing pulses may be applied to either tube.

Start-stop Multivibrator

The start-stop multivibrator is a widely used circuit, which never becomes free running and must always be triggered for each output pulse. This circuit is often called a one-shot multivibrator because it produces one cycle of a square wave each time it is triggered.

As shown in the schematic on page 312, the start-stop multivibrator consists essentially of a two-stage resistance-capacitance coupled amplifier with one tube cut off and the other conducting.

The balanced condition of the circuit is established by the arrangement for biasing the tubes. The grid of V_2 is connected to its cathode through the resistor R_2. The cathode of V_2 is connected to the cathode of V_1. As V_2 conducts, the voltage drop across R_k biases V_1 to cutoff. If V_2 is not conducting, V_1 cannot be cut off by the self-bias developed across R_k.



Action in Circuit When V_2 Conducts

The following are the steps in the action of the start-stop multivibrator which produce the waveshapes shown to the right.

From time t_o-t_1, V_2 is conducting heavily because its grid is at cathode potential and V_1 is cut off by the voltage drop produced across R_k by the plate current of V_2. Therefore the plate voltage of V_1 (e_{p_1}) is at B+ and the plate voltage of V_2 (e_{p_2}) is considerably less than B+.

At time t_1, a positive trigger pulse, sufficient in amplitude to raise the grid of V_1 above cut-off voltage, is impressed on the grid of V_1 through C_1. V_1 begins to conduct and the voltage at its plate decreases. This decrease passes through C_2, as the voltage across a capacitor cannot be changed instantaneously, and appears across R_2 as a negative-going voltage on the grid of V_2. This negative-going voltage on the grid of V_2 decreases the current through V_2 and the voltage drop across R_k decreases. This decrease in bias allows more current to flow in V_1 and further decreases the plate voltage of V_1 driving the grid of V_2 still more negative. This action is cumulative until V_2 is cut off and V_1 starts conducting. The switch from V_1 to V_2 is practically instantaneous as you can see from the waveshapes.

The circuit remains with V_1 conducting and V_2 cut off during the interval from t_1 to t_2 while C_2 discharges sufficiently toward the lowered value of e_{p_1} to allow the grid of V_2 to rise from its lowest value to cut-off voltage. At time t_2, V_2 begins to conduct the plate current of V_2 flowing through R_k, raises the cathode voltage of V_1 and causes its plate current to become less. The decreased plate current of V_1 causes the plate voltage of V_1 to increase sharply. This sudden rise in e_{p_1} is coupled to the grid of V_2, increasing its plate current still further. This flow of current through R_k biases V_1 beyond cutoff. The plate voltage of V_1 tends to rise to B+ instantaneously, but C_2 is being charged at the same time as shown from time t_2 to t_3.

The circuit has now come back to its original balanced state and will remain so until another positive pulse arrives and causes V_1 to conduct at time t_4.

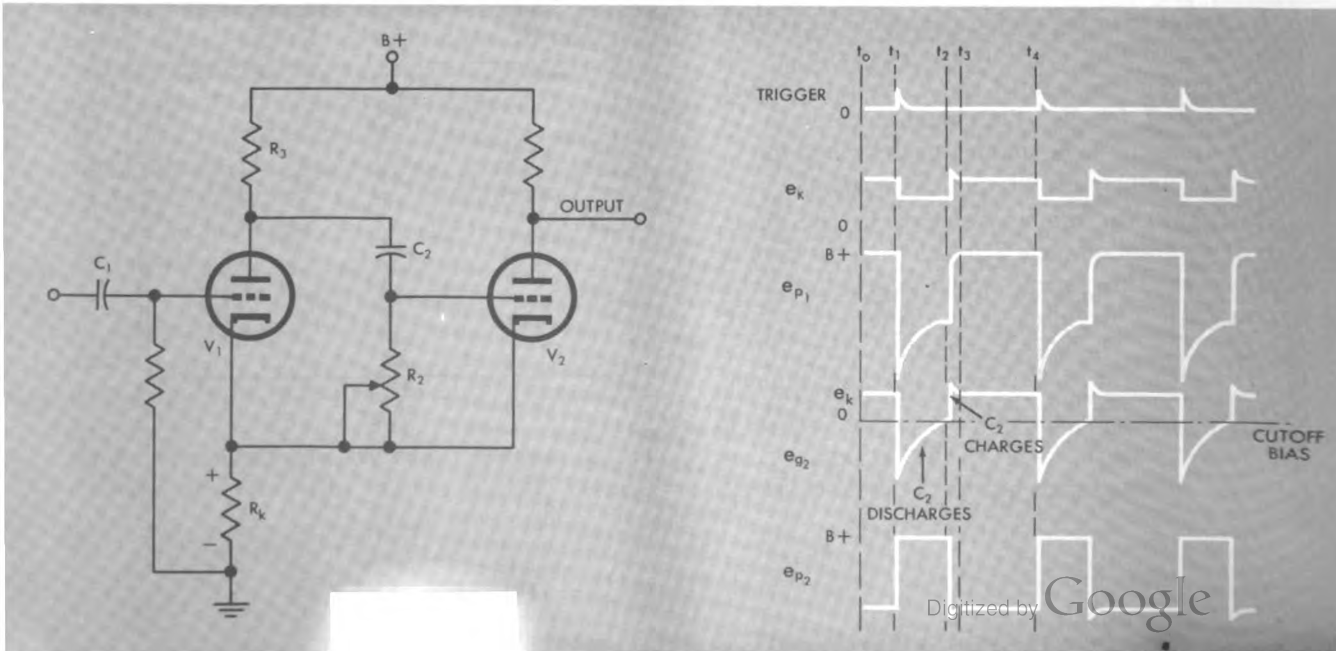
Note that every positive trigger pulse which causes V_1 to conduct results in a large positive-going pulse from the plate of V_2 and a negative-going pulse from the plate of V_1. The length of the output pulses produced is controlled by the time constant C_2R_2. If the resistance of R_2 is increased, the length of the output pulses is increased. An output pulse is produced for each input trigger pulse.

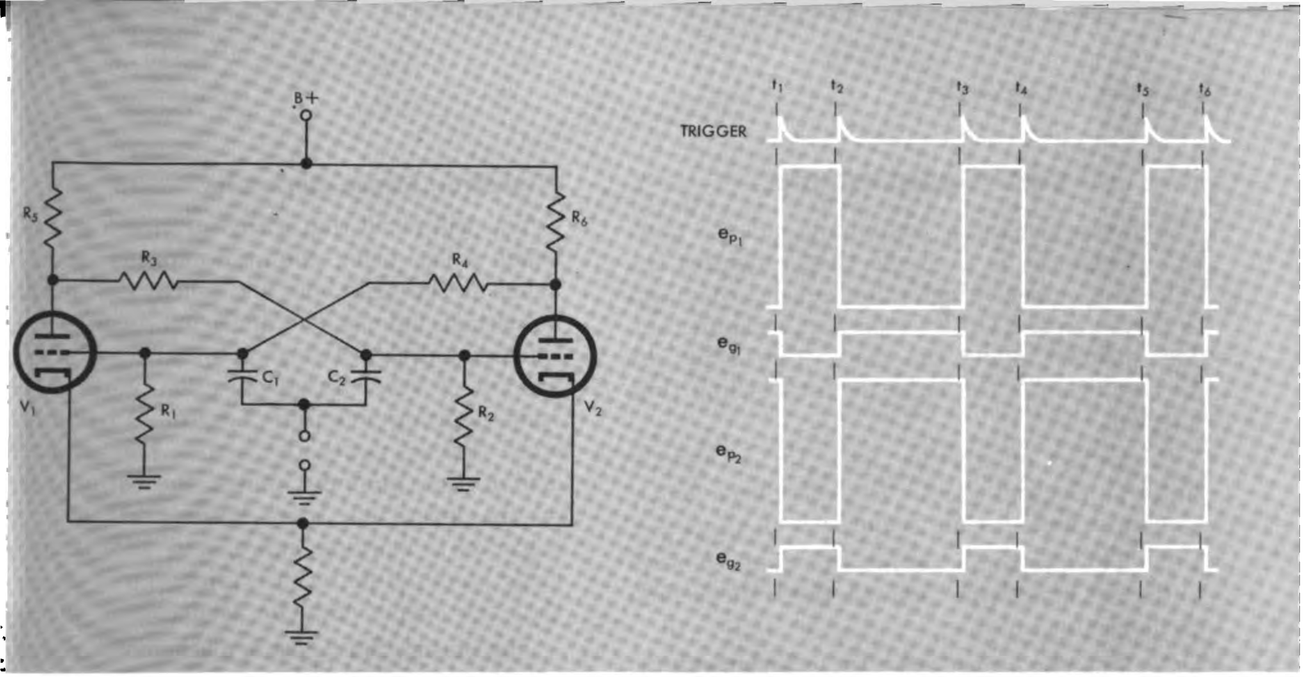
ECCLES-JORDAN MULTIVIBRATOR

The Eccles-Jordan multivibrator shown in the illustration on the next page is a very stable circuit; therefore, its output pulses can be used to gate circuits where a high degree of accuracy is required.

When power is applied to the circuit and no trigger pulses are applied, the following action takes place until stability is reached. As current begins to flow, the plate voltage of each tube will begin to drop. Since the circuits cannot be

Start-stop Multivibrator and Waveshapes





Eccles-Jordan Multivibrator and Waveshapes

perfectly balanced, the plate voltage of one will drop faster than the other. Assume that the plate voltage of V_2 drops faster and causes the control grid voltage of V_1 to drop faster than the control grid voltage of V_2. This faster drop of voltage on the control grid of V_1 will cause its plate voltage to remain high. Since the plate voltage of V_1 determines the control grid voltage of V_2, V_2 conducts more heavily and causes its plate voltage to drop even more. This cumulative action continues until V_2 conducts very heavily and its plate voltage drops to a low value. This drop drives the control grid voltage of V_1 down. The combined effect of the drop in control grid voltage and the cathode voltage developed by the conduction of V_2 is to cut off V_1. The circuit remains in this condition until a trigger pulse is applied.

This condition can just as easily be reversed if the plate voltage of V_1 drops faster than the plate voltage of V_2. Therefore, the Eccles-Jordan has two conditions of stability, and, consequently, two trigger pulses are needed to produce one complete cycle of operation. The trigger pulses may be applied to the grids, cathodes, or even the plates of the tubes and may be either positive or negative in polarity. The time between trigger pulses determines the time duration (width) of the resulting rectangular wave produced by the multivibrator.

With V_2 conducting and V_1 cut off as described, if positive trigger pulses are applied as shown in the illustration above, the following events take place within the circuit producing the waveshapes shown.

The first positive pulse applied to the grids of both tubes at time t_1 brings V_1 into conduction. The pulse has no effect on V_2 because V_2 is already conducting very heavily. As plate current increases through V_1, less voltage is dropped across R_2. Therefore the grid of V_2 becomes less positive. This decrease in the grid potential of V_2 causes an increase in the plate potential of V_2 and causes the grid of V_1 to become more positive. This cumulative action continues until V_2 is cut off and V_1 is conducting heavily.

The circuit will remain in this condition until another trigger pulse is applied to the grid at time t_2. This positive pulse brings V_2 into conduction. As plate current increases through V_2, less voltage is dropped across R_1. Therefore the grid of V_1 becomes less positive. This decrease in grid potential of V_1 causes an increase in the plate potential of V_1 and the grid of V_2 to become more positive. This cumulative action continues until V_1 is cut off and V_2 is conducting heavily.

An output may be taken from the plate of either tube. Both waveshapes will have exactly the same duration. Since the circuit will produce a square wave only when it is triggered by two trigger pulses, both the output frequency and the width of the output waveshape are controlled by the frequency of the trigger pulses.

SWEEP GENERATORS

A sweep generator, commonly called a *time-base* generator, is a circuit that generates the voltage or current which causes an electron beam to move across a CRT screen at a constant rate. The distance along the trace formed by the moving beam is the basis for establishing time intervals.

Regardless of whether the CRT is electrostatic or electromagnetic, the time-base generator must meet certain requirements:

1. The electron beam must be made to move so that its image (spot) on the screen moves in a predetermined path.
2. The spot must move at a linear rate so that range measurements along its trace are linear. For example, if one inch of time base equals 5 miles, then two inches are equal to 10 miles anywhere along its length.
3. At the end of its trace, the spot must return rapidly to its point of origin; this return is known as *flyback*.
4. To be effective in giving distance, the exact duration of the trace must be known.
5. The flyback time must be as short as possible.

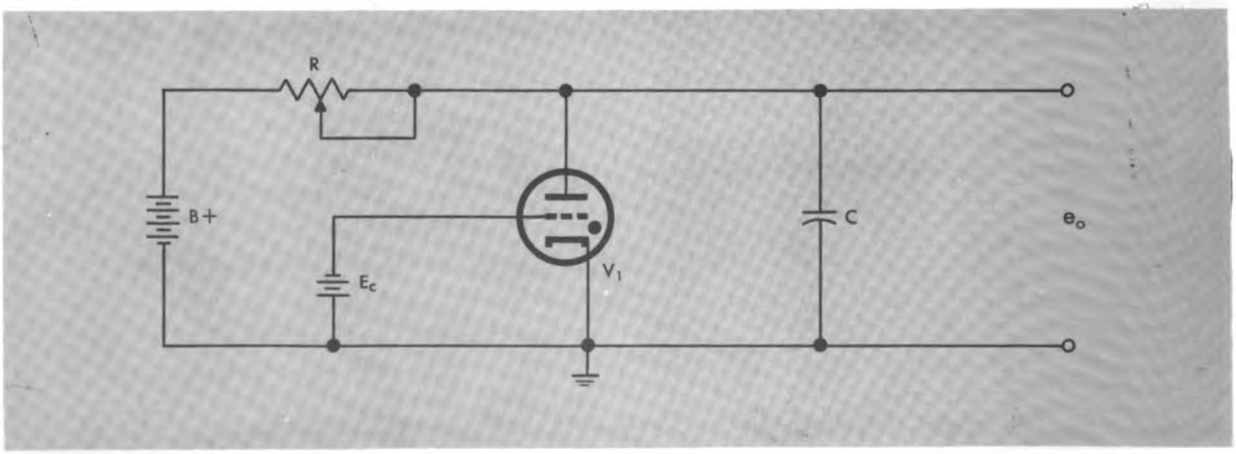
To fulfill the above requirements in an electrostatic CRT, a sawtooth voltage may be applied to the deflection plates. In an electromagnetic CRT, a sawtooth wave of current must pass through the deflection coils. This chapter is devoted to the circuits which produce sawtooth waveshapes.

Thyratron Sweep Generator

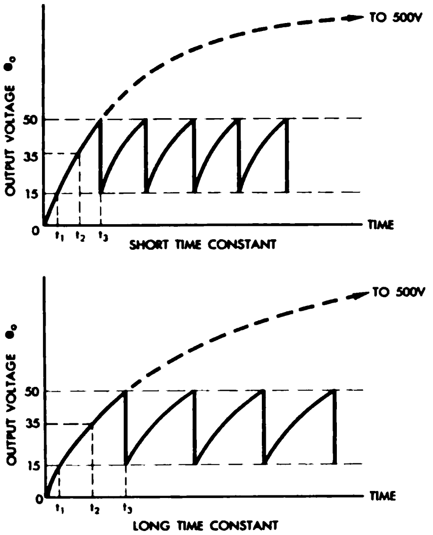
You will recall that a thyratron is a gas-filled triode whose firing potential depends upon the value of the grid bias and whose deionization potential is independent of all circuit components (including bias). Using a thyratron makes it possible to lower the deionization potential and vary the firing potential of gas tubes to obtain a more linear sawtooth wave. Lowering the deionization potential makes the tube operate on the lower and more linear portion of the capacitor's charging curve, and adjusting the firing potential limits the variation in voltage so that the output may be as linear as possible.

The illustration below is a schematic diagram of a thyratron sawtooth generator. For purposes of explanation, let us assume that V_1 is a thyratron whose firing potential is ten times its grid bias and whose deionization potential is 15 volts. For example, if the bias (E_c) is 5 volts, the ionization potential is 50 volts. For good linearity of the sweep, the sawtooth voltage should lie on the first 10 percent of the capacitor's charge curve. Thus, if B+ is 500 volts, the charge of the capacitor should be limited to 50 volts. To obtain a firing potential of 50 volts, a bias of 5 volts is required. When these constants are used in the circuit (B+ = 500 v, $E_c = -5$v), a sawtooth output voltage is obtained.

The action of the circuit is as follows: The capacitor charges at a rate determined by values of R and C until the firing potential of the thyratron is reached. The capacitor then discharges through the tube until the extinction



Thyratron Sawtooth Generator



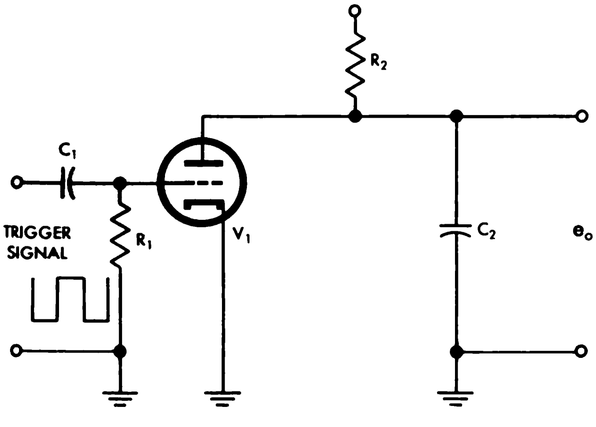
Output Waveshape of a Thyratron Sawtooth Generator

potential of the tube is reached. The capacitor recharges, and the process continues.

Since the time constant (R×C) determines the time it takes for the voltage on the capacitor to reach the firing potential, it obviously controls the frequency of the generated sawtooth wave. In the illustration above, the effects of a short time constant and of a long time constant on the output waveshape are shown. From the illustration you can see that, as the time constant increases, the frequency decreases. This conclusion may be reached if we recall that

$$f = \frac{1}{T},$$

where f is the frequency of a wave and T is the period of a wave. Since an increase in time constant increases the time (period) of each wave, the frequency of the wave decreases. Note, however, that neither the amplitude nor the linearity of the sawtooth wave is influenced by a change in time constant.



Hard-tube Sweep Generator

The frequency of a thyratron sawtooth generator is usually capable of meeting all the requirements of time-base generators when it is used for cathode-ray tubes in test oscilloscopes, but in radar the requirements of the time base make the use of the thyratron sawtooth sweep impractical.

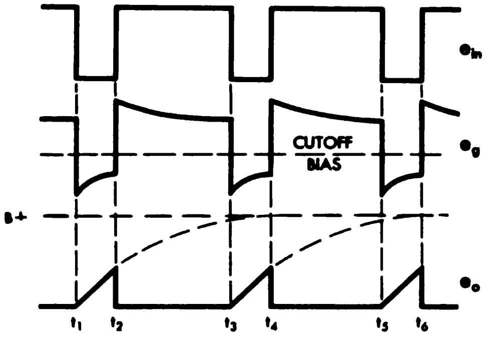
The Hard-Tube Sweep Generator

To obtain a sweep voltage that fulfills the requirements of the radar time base, a hard-tube sweep generator is used. This type of generator is so named because a vacuum (hard) tube replaces the gas (soft) tube of the thyratron sawtooth generator. The major points of difference in the operation of the two types of sweep generators are outlined in the table.

THYRATRON	HARD TUBE
TIME OF SWEEP	
Depends upon circuit constants: E_b, E_c, R, and C.	Depends upon time of the external signal.
SWEEP VOLTAGES	
Produced continuously as soon as circuit is connected.	Produced only when an external signal is applied.

Note that an external signal is required to operate the hard-tube sweep generator. This signal is called a trigger signal because it starts or "triggers" the circuit into operation. As shown in the illustration at lower left, the trigger signal is applied to the grid of a triode with zero bias. This input wave is usually the output of a square-wave generator.

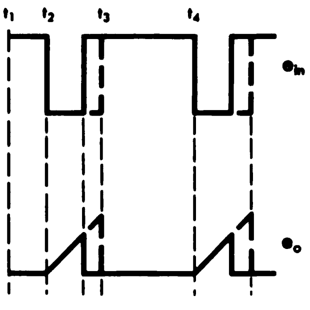
Let us examine the circuit. When no signal is applied to the grid of V_1, the tube conducts heavily and is a short circuit across C_2. The voltage across C_2 is equal to the voltage across V_1, since they are in parallel. If the grid of V_1 is made sufficiently negative, the tube ceases to conduct and V_1 acts as an open circuit. With V_1 open-circuited, a series RC circuit consisting of B+, R_2, and C_2 is formed. As a result, C_2 begins to charge to B+. The time required for C_2 to charge to B+ depends on the time constant (R_2C_2) of the circuit.



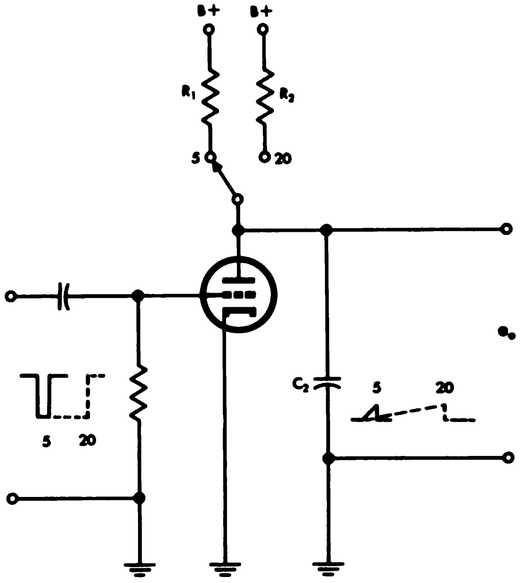
Waveforms of Hard-Tube Sweep Generator

With the aid of the waveshapes shown at top of this page, let us see the effect of applying a negative rectangular pulse to the grid of V_1. As soon as the negative rectangular wave is applied to the grid of V_1, the tube is cut off. C_2 immediately starts to charge toward B+. If V_1 were cut off long enough, C_2 would charge to B+ as indicated by the dotted line. Usually, however, V_1 is cut off for a relatively short period of time and C_2 charges as long as the trigger pulse is applied to the grid of V_1, as shown in the illustration on page 315. As soon as the trigger pulse ends, V_1 immediately starts conducting heavily, forming a short circuit through which C_2 discharges very rapidly. C_2 at this time will have a very low potential across it, equal to the voltage across V_1. The circuit will remain in this condition until another trigger pulse is applied.

As you can see, the duration of the sweep is determined by the time interval t_1 to t_2. If the same circuit is to be used for longer sweeps, the interval from t_1 to t_2 must be increased. As you can see by the dotted lines in the illustration below, a wider trigger pulse allows C_2 to charge to a higher value of voltage. The voltage of the sweep must be the same for all ranges if the same CRT is used. To prevent an increase in voltage, the RC time constant must be made greater so that the capacitor will charge at a slower rate. This may be accomplished by changing the plate load resistor in the circuit to a larger one as shown in the illustration below. If R_1 were left in the circuit for both ranges, the sweep would extend off the edge of the CRT on the 20-mile range, and that portion of the sweep would be of no value to us. By switching R_2 (a larger resistor) in the circuit when we increase the range to 20 miles, C_2 charges at a slower rate and the amplitude of the sweep is now the same for both ranges. The sweep will now extend just to the edge of the CRT for both ranges.



Wider Trigger Pulse Allows C to Charge to Higher Value



Increasing Range of Sweep

LIMITERS OR CLIPPERS

Limiting circuits are circuits which remove all or part of either the positive or negative portion of an input wave. Circuits which perform this function are called ***limiters*** or ***clippers***. Limiters are useful in a variety of ways. They are applicable in waveshaping circuits where the extremities of the input signal are to be squared off. A sine wave can be converted into a rectangular wave by a limiter circuit. A peaked wave may be applied to a limiter to eliminate either the positive or the negative peaks from the output. Limiters are also used to prevent a voltage from swinging too far in either the positive or negative direction.

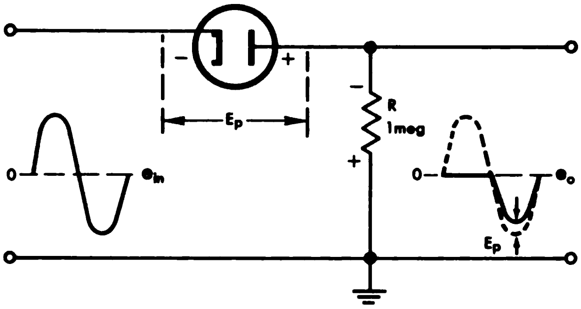
Series Diode Limiters

Diodes are very useful for limiting since they conduct current only when the plate is positive with respect to the cathode. The illustration to the right shows a series-connected diode (one in which load is connected in series with the tube) that is used to limit the positive input cycle of a sine wave. The input voltage is e_i and the output is e_o. When any input voltage is applied to the input terminals of the diode, the output follows the input only when the negative input cycle is applied.

When the positive cycle is applied, the cathode becomes positive with respect to the plate and consequently the tube cannot conduct current. Since there is no current flow, there is no voltage developed across the 1-megohm output resistor. Therefore, the positive input cycle is limited to zero in the output as shown by the solid line in the output. The dotted portion is the clipped part of the wave.

When the negative input cycle is applied, the plate becomes positive with respect to the cathode and the tube conducts. The current which flows through the tube develops a voltage across the internal resistance of the tube and the load resistor. These two resistances act as a voltage divider and divide the applied voltage as shown by the dotted and solid portions in the negative portion of the output wave. The voltage dropped across the resistance of the tube is represented by e_p. Except for the small drop across the tube, the output is nearly equal to the negative input. Thus, the diode has limited or clipped the positive input cycle, and has virtually reproduced the negative input cycle both in shape and magnitude.

When the diode is reversed as shown in the illustration to the right, the diode limits the negative alternation of the input sine wave because during that alternation the plate is negative with respect to the cathode. During the positive alternations, the tube conducts because its plate is positive with respect to the cathode. The current through the diode develops a voltage across the resistor which follows the positive cycle of the input and is equal to it except for the small drop across the tube itself. The output wave is shown by the solid line and the part that is clipped by the dotted line.

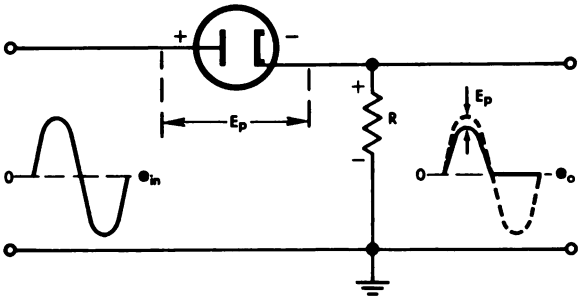


Series Diode Limiter Used to Clip Positive Voltage

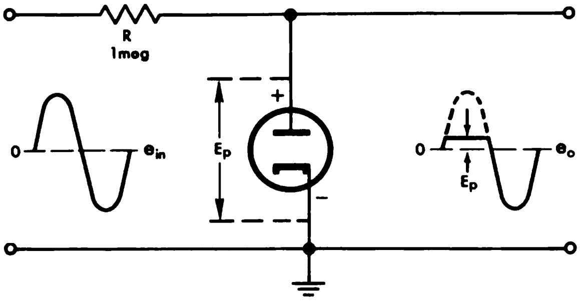
Shunt Diode Limiters

Another method of using a diode as a limiter is to connect it in parallel with the load. In the schematic shown at the top of page 318, the diode is connected to limit the positive input cycle. When the positive cycle of the input voltage is applied to the plate the diode conducts. Therefore, the current flows through a voltage divider consisting of the 1-megohm resistor and the internal plate resistance. This voltage divider divides the applied voltage with most of the voltage dropped across the 1-megohm resistor as shown by the dotted portion of the waveshape. The remainder, shown as a solid line, is dropped across the tube. Since the output terminals are across the tube, the output is very small during the time when the input cycle is positive. The output is limited or clipped to practically zero voltage during this time.

During the negative portion of the input cycle, the plate is negative with respect to the cathode and current does not flow. Consequently, there is no voltage drop across the resistor. However, since the tube appears as an open circuit, the applied voltage appears across its output terminals. Thus, the output voltage follows and is equal to the input voltage during the time the negative cycle is applied to the tube.



Series Diode Limiter Used to Clip Negative Voltage



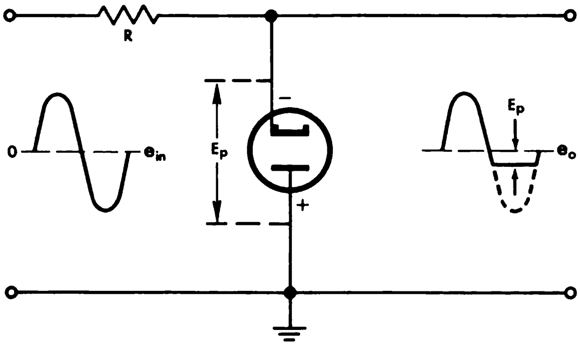
Shunt Diode Limiter Used to Clip Positive Voltage

The negative input cycle may be removed by connecting the tube as shown in the illustration below. When connected in this manner, the tube conducts when the cathode is negative and virtually all the applied voltage appears across the resistor and a very small amount across the tube. Therefore, since the output is taken from across the tube, the negative input cycle has been practically reduced or limited to zero.

Biased Limiters

The input voltage can be limited to some value other than zero by maintaining the plate or cathode at that voltage by means of a battery or a biasing resistor. Each of the limiting circuits in the illustration at the top of the next page employs a battery to supply a biasing voltage.

Circuit A is designed to limit the swing of the positive input cycle to +50 volts. The input voltage (e_i) for these circuits is 115 volts with a peak value of 163 volts. The battery connected in the cathode circuit maintains the cathode 50 volts positive with respect to the plate before the input is applied. As long as the input voltage is less than 50 volts, the tube cannot conduct; however, as soon as the input exceeds 50 volts, current flows. The low-resistance of the tube while current flows effectively connects the upper output terminal of the circuit to the positive terminal of the battery. Therefore, during the portion of the positive input cycle when the voltage exceeds the 50 volts on the cathode, the output voltage equals the battery voltage of 50 volts. The difference between the input voltage and this voltage (neglecting the small voltage drop across the internal resistance of the tube) appears across the 1 megohm resistor.



Shunt Diode Limiter Used to Clip Negative Voltage

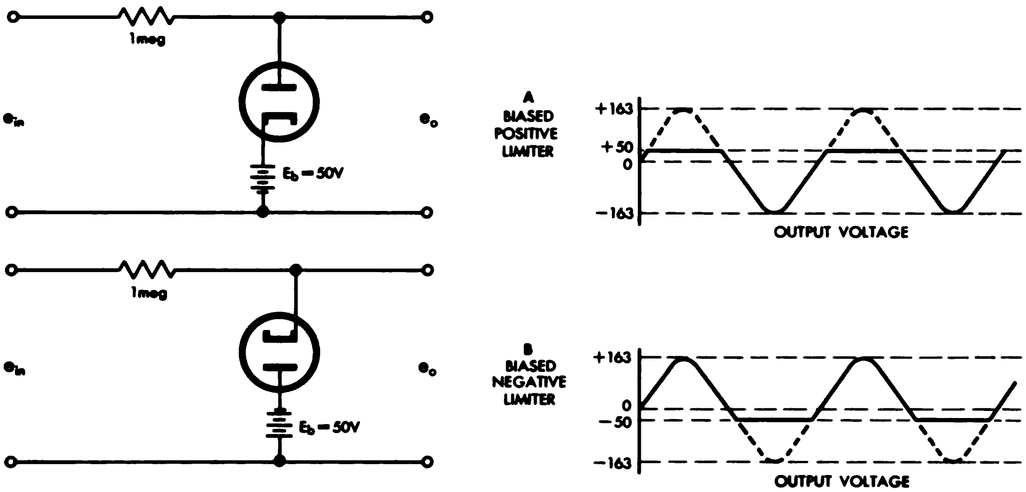
In diagram B, the battery E_b in the diode plate circuit makes the plate 50 volts negative with respect to the cathode. As long as the input is positive or less negative than E_b the diode acts like an open circuit, and the voltage across the output is equal to the input. When the input becomes more negative than E_b, the cathode is negative with respect to the plate and the diode conducts. This conduction effectively connects the upper output terminal to the negative terminal of the battery. Therefore, during this part of the input cycle, 50 volts appears in the output. The rest of the 163 volts input, neglecting the small drop across the tube, appears across the 1-megohm resistor.

Shunt diodes may be used also to limit the amount to which the input voltage can drop. In other words, only the peaks of waveform are reproduced in the output. In circuit A of the lower illustration on page 319, the diode conducts as long as the voltage of the input waveform is below the positive voltage of the battery. The output voltage varies between the positive level of the battery voltage and the positive extremity of the input waveform as shown.

In circuit B, the entire portion of the input waveform above the negative potential of the battery causes the diode to conduct. This produces an output voltage which varies between the negative level of E_b and the negative extremity of the input as shown. In both cases, the difference in voltage between the value of E_b and the applied voltage while the diode conducts is dropped by the 1-megohm resistor.

Crystal Diode Limiter

Although most radar circuits use diode vacuum tubes in limiter circuits, the crystal diode or crystal rectifier is becoming increasingly popular for limiting. One type of crystal diode



Biased Limiters

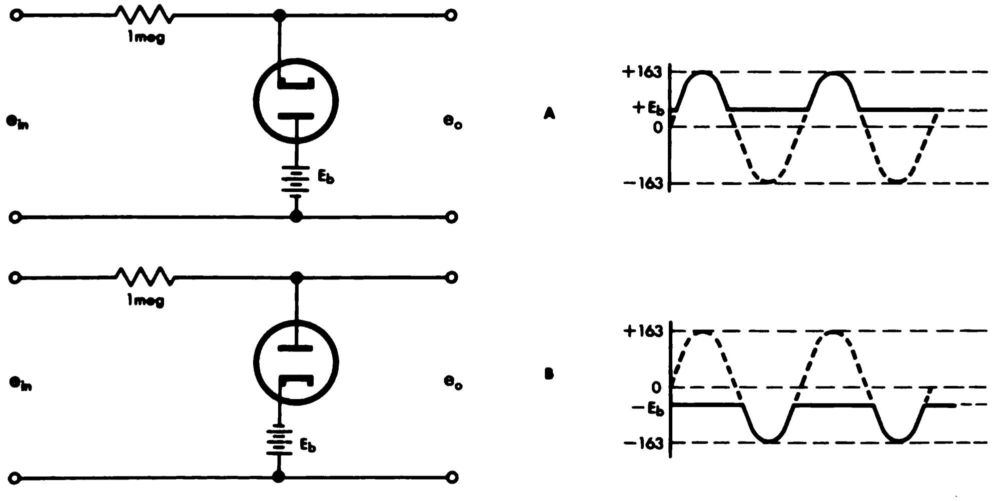
consists of a small piece of silicon and a pointed tungsten wire contactor. When an AC voltage is applied to the input terminals of the crystal, little current passes from the crystal to the wire because of high resistance in that direction, but current flows readily from the wire to the crystal. Since, like a diode vacuum tube, it conducts more current in one direction than in the other, the crystal assembly often is called a crystal diode. Not only is the crystal almost as efficient as the vacuum tube, but it requires no heater voltage and it is entirely free of interference from the power frequency. It may be used in any of the diode limiter circuits previously described. The crystal limiter illustrated at the top of page 320 limits or clips the positive cycle of the input waveform as shown.

Double Diode Limiting

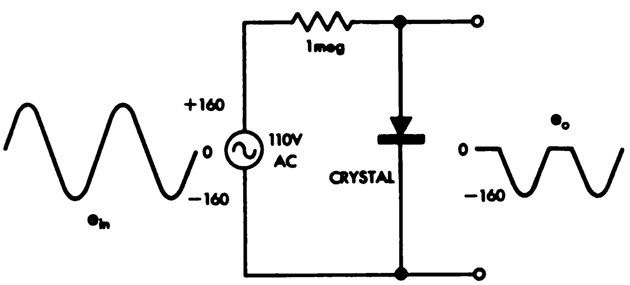
When a pair of biased diodes are connected as shown in the lower illustration on page 320, both the negative and positive half-cycles of the input voltage can be limited, leaving a waveshape that is almost square. The voltage at which limiting occurs in each direction may be varied by varying the bias voltage on the diode. In the circuit shown, tube V_1 conducts and limits the positive half-cycle, and V_2 conducts and limits the negative half-cycle.

Triode Limiters

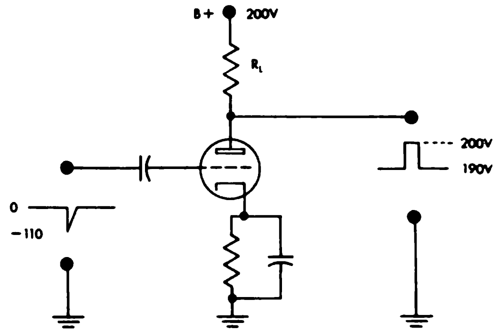
Another type of limiter commonly employed makes use of the triode tube, taking advantage of the fact that the tube can be driven to cutoff. Current through a vacuum tube can flow



Reproducing Peaks of Waveforms



Positive Crystal Limiter



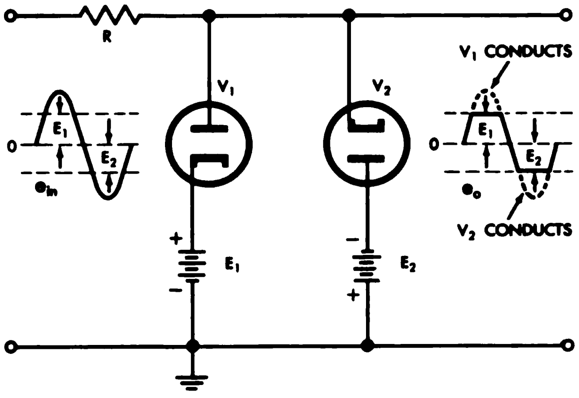
Cutoff Type Triode Limiter

only from cathode to plate; therefore, plate current cannot become a negative value. When the grid is driven to cutoff, the plate current is decreased to zero and remains at zero during the time the grid is below cutoff. Since no current flows through the plate circuit when the tube is cut off, there is no voltage developed across the load resistance and the plate is maintained at the full value of the plate supply voltage. Thus, a type of limiting is achieved in which the positive part of the plate waveform is flattened or clipped as a result of driving the grid beyond cutoff.

The illustration at the right above shows the operation of the cutoff type triode limiter.

Assume that cutoff for the tube is −10 volts and the stage has a gain of one:

A negative spike varying from zero to −100 volts applied to the circuit will cause the tube to conduct for only 10 volts of the input and then stop. Thus, current will flow through R_L only for the portion of the input signal that permitted the tube to conduct, and the output will be limited to 10 volts. The output is shown as varying from +190 to +200 volts. If the stage had a gain of two, the amplitude of the output would be 20 volts instead of 10.



Double Diode Limiter

CLAMPING CIRCUITS

Whenever a given waveshape is applied to an RC coupling circuit, the output waveshape will have an axis or reference line established. The wave becomes AC in nature and varies above and below the reference line. In certain applications, this is undesirable. For example, if a sawtooth wave is to be used as a sweep voltage for a certain cathode-ray tube, the change in amplitude should vary in one direction only. This can be accomplished by clamping circuits.

A *direct current restorer* (DCR) is a circuit that shifts a waveshape so that it is all above or all below a certain voltage—often zero voltage. It is also called a *clamping circuit*, or *clamper*, because it clamps the top or bottom of a voltage waveshape to a given reference voltage.

Negative Clamper

A simple clamping circuit is shown at A in the illustration on the next page. The input voltage is a 100-volt peak-to-peak square wave varying from plus 50 volts to minus 50 volts. When the first positive half-cycle is applied, C_1 charges through V_1 to the full 50 volts. The output voltage is zero during this half-cycle because V_1 is a short-circuit across R_1 while V_1 conducts.

During the negative half-cycle, a negative 50 volts is applied to the circuit. C_1, which has been charged to 50 volts during the positive alternation, now begins to discharge through R_1.

Both the applied voltage and the C_1 voltage cause current to flow through R_1 in the same direction. The diode does not conduct during this cycle because its plate is negative. Capacitor C_1 discharges very little because R_1C_1 is a long time constant. Effectively, we have two 50-volt sources of potential in series causing a 100-volt drop across R_1. At B in the illustration is an equivalent circuit of the negative clamper during this negative half-cycle, with a 50-volt battery replacing the voltage across capacitor C_1. The output across R_1 will then be minus 100 volts in amplitude. Thus the clamper has shifted the waveshape from a half negative and half positive voltage in the input to an output voltage which is zero during the positive alternation and 100 volts negative during the negative alternation. This circuit is, therefore, called a *negative* clamper.

Positive Clamper

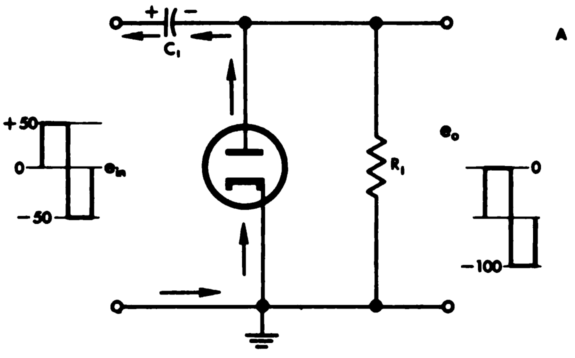
To clamp a waveshape in a positive direction, simply reverse the diode as shown in the last illustration at the right. Here the tube conducts during the negative alternation of the input wave and does not conduct during the positive alternation. Note that the capacitor charges on the negative alternation with the polarity shown. The output waveshape is therefore clamped from zero to positive 100 volts as shown.

Note that the value of R must be very large to accomplish clamping. The capacitor of the DCR charges through a short time constant (consisting of the capacitor and the resistance of the tube) while the tube conducts and discharges through a long time constant (consisting of R and C) while the tube does not conduct. The capacitor therefore discharges very little during the period of nonconduction and may be considered to be a battery with a constant charge as shown in the illustration at the right.

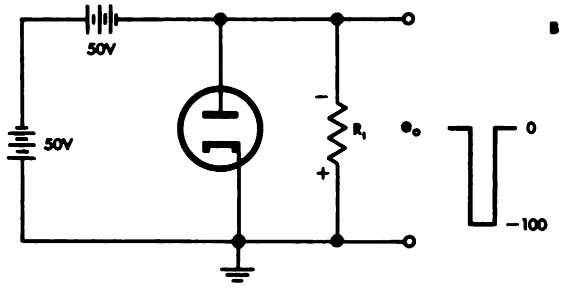
Although all the DCR's discussed were shown with square-wave inputs, the shape of the input wave does not matter. A clamper will clamp a wave to a given reference level regardless of whether a sine wave, sawtooth wave, or square wave is applied. The effect of positive clamping on a sine wave is shown in the illustration at the top of page 322.

Biased Restorers

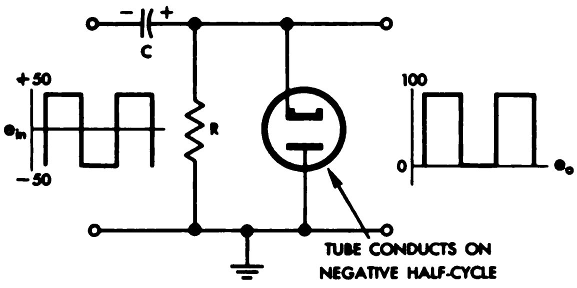
A biased restorer clamps a waveshape at a reference level other than zero by inserting a DC voltage equal to, and of the same polarity as, the voltage at which clamping is desired. The only change in the circuit from the positive and negative restorers discussed previously is the insertion of the DC potential in series with the diode. The circuit at A in the illustration at the bottom of page 322 is a positive clamping circuit that shifts the entire waveshape in a positive direction. The DC source inserts an additional positive 50 volts, which shifts the entire waveshape to 50 volts above zero. Thus, the waveshape which originally varied from minus 100 volts to plus 100 volts, now varies from plus 50 volts to plus 250 volts.



Negative Clamper



Equivalent Circuit of Negative Clamper during Negative Alternation

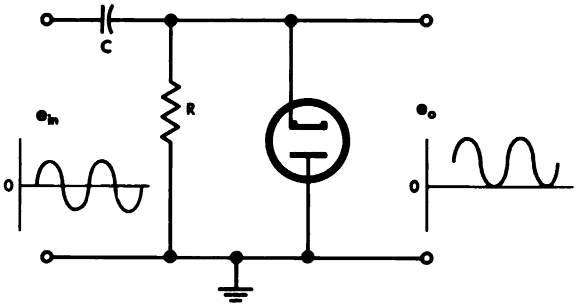


Positive Clamper

The electrical operation may be explained as follows:

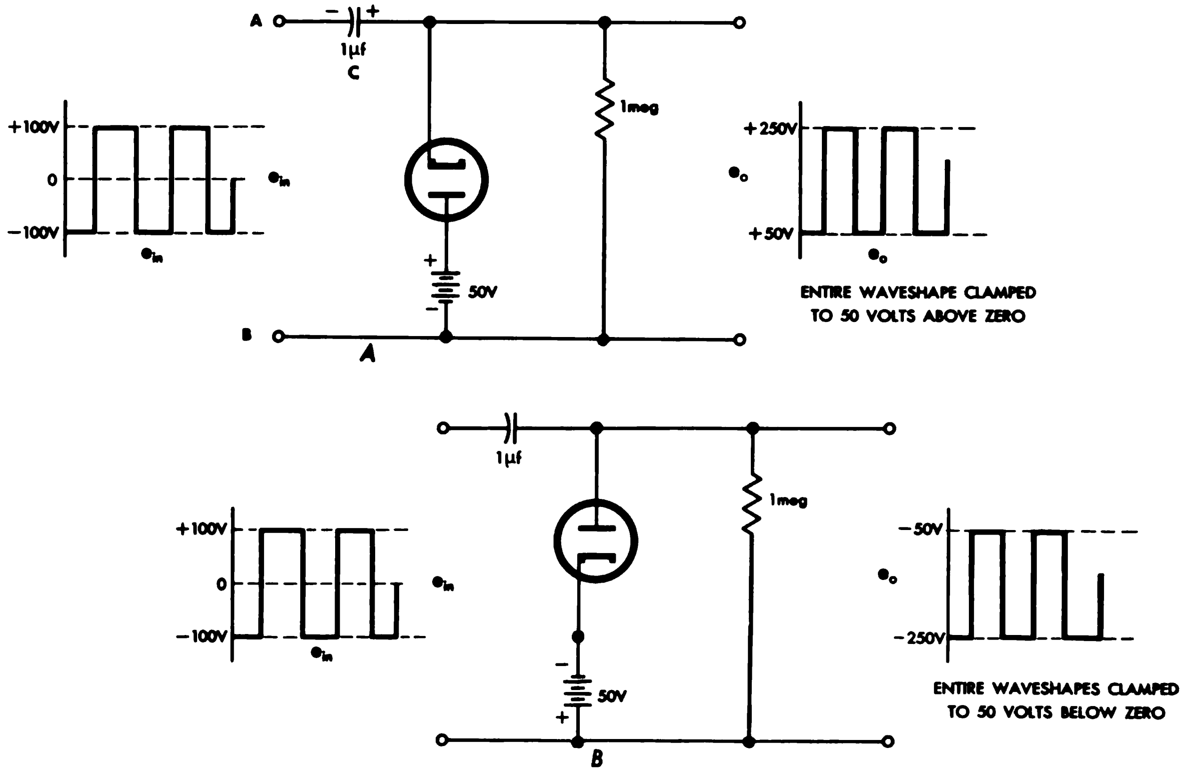
1. When the negative alternation is impressed between points A and B, A becomes negative with respect to B. The cathode of the diode is then negative with respect to the plate and the tube conducts charging C to 150v (—100v plus the +50v of the battery). When this occurs, the +50v of the battery is felt across R (the output) because a complete circuit exists from the negative terminal of the battery through R and the tube to the positive terminal of the battery. This yields an output of +50v when the input is —100v.

2. When the positive alternation of the input is impressed between points A and B, A becomes positive with respect to B. A potential of +100v then exists between A and B. This voltage plus the +150v across the capacitor is then applied across R because a complete circuit exists through R. This becomes the +250v output. The battery voltage cannot now appear across R (and hence does not appear in the output) because there is not now a complete circuit from the negative terminal of the battery through R and the tube (the tube cannot conduct) to the positive terminal of the battery. The tube cannot conduct during this period because the cathode is +200v with respect to the plate.



Effect of Positive Clamping on a Sine Wave

The circuit at B is a negative clamping circuit that shifts the entire waveshape in a negative direction. Since the circuit is biased at —50 volts, the output waveshape varies in a negative direction from —50 volts to —250 volts.



Biased Restorers

Basic Radar Principles

If you are old enough to remember much about the last war, you may recall the mystery, the cloak-and-dagger atmosphere, which surrounded radar. That secrecy came about not because radar was particularly complicated but because it had been classified as a "secret weapon." Actually, radar circuits are no more complicated than many other electronic circuits, and even the newer radar sets are only refinements of combinations of circuits which appeared in early sets.

Essentially, radar is a radio device used to locate airplanes, ships, or other objects in darkness, fog, and storms, and its kinship with radio is even emphasized by the derivation of the word radar.

The word *radar* is a contraction of the expression *RA*dio *D*etection *A*nd *R*anging. The italicized capitalized letters, as you can see, form the word radar. Radar is an application of radio principles to detect objects that cannot be observed visually and to determine their direction, range, and elevation.

PRINCIPLES OF RADAR OPERATION

Just as you need to know the principles of internal combustion to understand how a gasoline engine works, so also do you need to know the fundamental principles which radar uses to understand how it works. One of the basic physical laws which apply to radar is that time is required for an object to pass between two points in space. Similarly, it requires time for energy, say in the form of sound or light waves, to pass through a given distance. Sound waves, we know, travel at about 1,090 feet per second while light waves travel at about 186,000 miles per second. This means that an observer would be aware of an interval between the instant he sees a flash of lightning and the instant when he hears the sound of thunder. Or an even better example in that it covers less distance, is that of the caddy on the golf course a couple of hundred yards away from the tee. He watches the golfer swing at the ball, but there is a brief lapse before he hears the impact. The time interval naturally is proportional to the distance from the observer to the flash or the golfer.

Another example of how sound waves require time to travel is an echo, a form of wave reflection. If you shout toward a cliff, or some other sound-reflecting surface, you hear an echo from the direction of the cliff. What happens is that the sound waves generated by the shout travel through the air until they strike the cliff; they are then reflected so that some waves return to you. An interval of time passes between the instant you shout and the instant you hear the echo. Obviously, the farther you are from the reflecting surface, the longer you'll have to wait for the echo. Sound waves travel at 1,090 feet per second. So if you were 2,180 feet from a cliff when you shouted, it would take four seconds for the echo to get back to you — two seconds each way.

You know from experience that when you make a loud noise, such as a shout or a gunshot, you often hear an echo of that noise. Sometimes you hear several echoes from the same noise apparently coming from different directions. Without looking, you know that there is a building, hill, grove of trees, or some other object nearby in the direction of the first echo, and another similar object still farther away in the direction of the second echo. The loud noise consists of sound waves which start from your position and strike the object. Upon striking the object, some sound waves are reflected. As these reflected sound waves move by you, you hear them. Since sound travels at a constant rate, it takes a certain amount of time for the sound waves to travel to the object and return to you as an echo. And still more time is required if the object is farther away. Furthermore, you are able to tell the direction of the object which caused the echo because of your ability to hear. With this faculty, your brain tells you that the object creating the echo is to the right, left, or in front of you.

In the same way, radar depends on the principle of creating and picking up an echo — a radio echo. Radar sets send out short, strong bursts of radio energy. Many of these bursts of energy strike objects in the vicinity and are reflected back to the radar set. When the time required for the energy to go to the object and to return is carefully measured and translated into distance in miles or feet, it is possible to determine the distance of the object causing the echo. In early radar sets, two antennas were used to compare the strength of the energy reflected from two directions. From this comparison, it was possible to determine the direction from which the echo came. Modern sets employ various types of directional antennas to determine direction. Thus radar determines the distance and direction of an object creating a radio echo in the same way you might determine the direction and distance of an object causing a sound echo.

In radar terminology, the reflected energy is called the *radar echo* or simply the echo. The object which reflects the radio waves is called the *target* (principally due to military influence. Modern radar technique is highly developed. Sets are unbelievably accurate; in many cases they are almost completely automatic. Yet, regardless of how complicated they may seem, all depend upon this simple principle of creating and detecting a radio echo.

RADAR SYSTEMS

Essentially, a radar system consists of a transmitter which sends out radio signals, a receiver which is located at the same site, and an indicator which gives a visual indication of echoes returned by a target.

When a radio signal which has a constant frequency is emitted by a radio transmitter, radio waves, in a manner similar to light and sound waves, travel out in all directions and are reflected by any object that they strike. On striking the object, components of the wave are reflected and likewise travel in all directions. Some of the reflected waves return to the site of the transmitter originating them, where they are picked up by the receiver, providing it is tuned to the correct frequency.

This is a simple explanation. Naturally, there are some problems which must be solved. First, there is the problem of having a powerful transmitter near the receiver and operating at the same frequency. In order for the receiver to detect the reflected signal, and thus indicate the presence of the target, the signal from the transmitter must be prevented from affecting and blocking the receiver.

The second problem is measuring distance. If a continuous constant frequency signal were transmitted, it would be impossible for the receiver to distinguish between various echoes at different distances, for they would all be alike. Thus, some means must be provided to eliminate the problems involved in using constant frequency signals.

There are several systems which satisfactorily solve these problems. They are the frequency modulation system, the frequency shift system, and the pulse modulation system. The pulse-modulation system is the most important and most widely used in the radar sets of the present day. This system is therefore the only one discussed in this manual.

Pulse System

In the pulse system of detection, the transmitter is turned on for short periods and off for long periods. During the period when the transmitter is turned on, it transmits a short burst of radio energy called a *pulse*. When a pulse strikes any object, part of the reflected energy is returned to the receiver, where it is displayed on the screen of a cathode-ray tube which is capable of measuring periods of time as short as one-millionth of a second. Since the transmitter is turned off after each pulse, it does not interfere with the receiver between pulses.

To completely locate an object in space by radar pulses, the *range* (the distance of the target) and the *direction* (including the azimuth and elevation directions) must be determined.

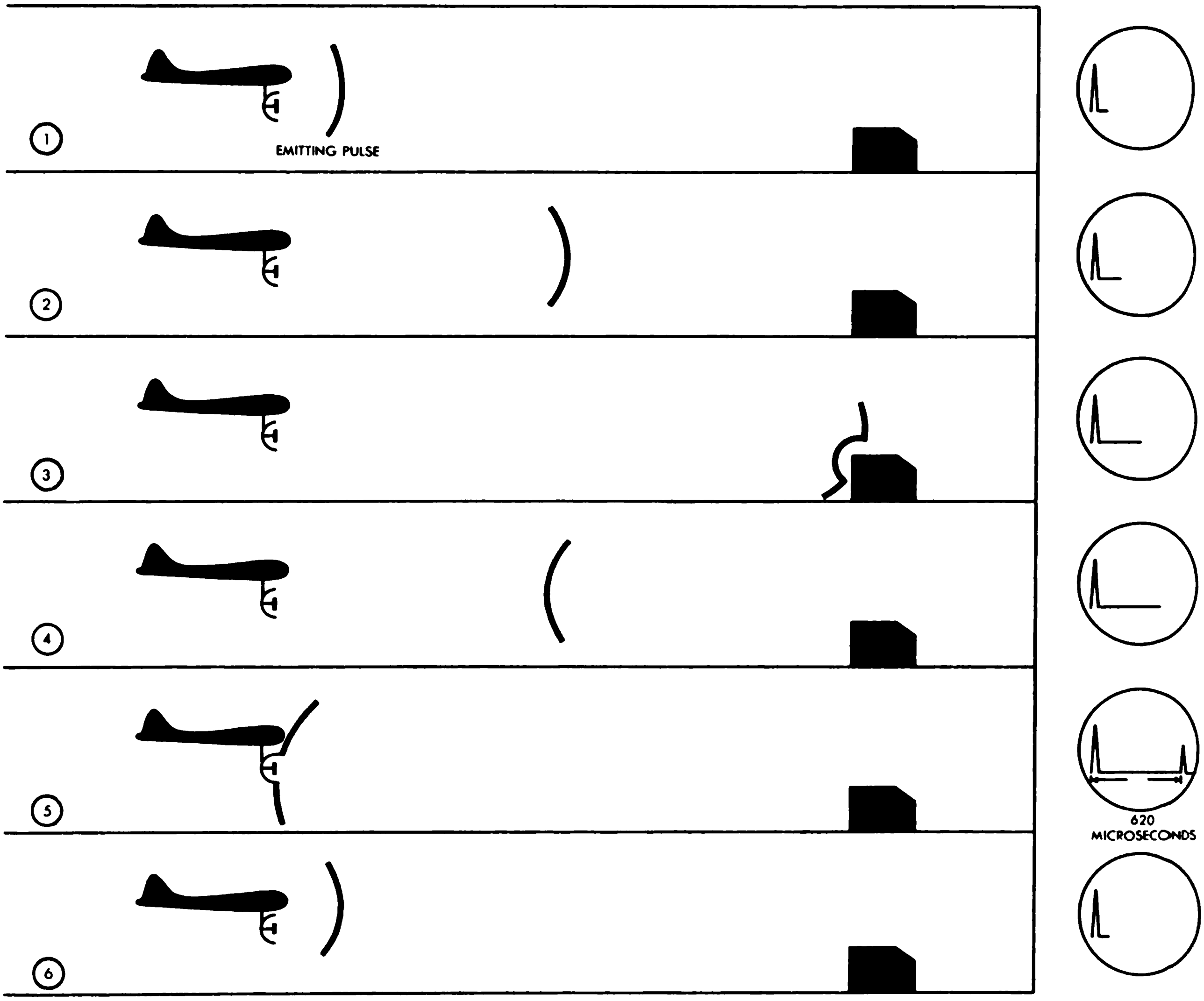
DETERMINING RANGE. The successful use of a pulse-modulated radar set depends primarily on its ability to measure distance in terms of time. When radio energy is radiated into space, it continues to travel with a constant velocity. Its velocity is that of light, or approximately 162,000 nautical miles per second. In more useful terms, radio waves travel a nautical mile in 6.2 microseconds. In radar applications, radio waves must travel twice the distance to a target (out and back). They therefore travel a *radar mile*, that is, a go and return mile, in 12.4 microseconds.

This constant velocity of radio waves is used in radar to determine the distance or range of a target by measuring the time required for a pulse to travel to a target and return. Suppose, for example, a pulse of radio energy is transmitted toward a target some distance away and the radar echo returns after 620 microseconds. Since energy travels a radar mile in 12.4 microseconds, the distance of the target is $\frac{620}{12.4}$ or 50 miles.

In order to use the time-range relationship, a time-measuring device must be used. The cathode ray tube is useful for this purpose since it responds to changes as rapid as one microsecond apart. A linear sweep which causes an electron beam to move across the screen of the cathode ray tube at a known rate of speed provides a time base.

The formation of the time base is shown in the illustration on the next page. In (1) a radar pulse is leaving the airplane. At the time the pulse is radiated, the spot on the screen of the cathode ray tube is deflected vertically for a brief instant, then it continues across the screen to the right. In (2) and (3) the pulse is traveling toward the target and the spot is moving across the screen. When the pulse strikes the target, the spot has moved halfway across the screen. In (4) the reflected pulse is returning. In (5) the reflected energy has returned to the receiver and there is a second vertical deflection of the spot on the right side of the cathode ray screen. The distance between the two upward deflections serves as the basis for determining the range of the target from the radar antenna. Assume, for example, that the set is designed so the spot moves across the cathode ray tube in 700 microseconds. The spot was almost to the end of the screen before the echo arrived, its position indicating that it took 620 microseconds to reach that point. Since radar waves travel one radar mile in 12.4 microseconds, the range of the target in the illustration is $\frac{620}{12.4}$ or 50 miles away. The last illustration shows another pulse being emitted and the start of the formation of a new time base.

DETERMINING DIRECTION. Two dimensions must be considered in determining the direction of a target. One is azimuth, which is the relative horizontal direction of the target with respect to some direction reference expressed in degrees. For example, this direction may be expressed with reference to true north or with reference to the heading of the airplane. The other dimension is elevation, which, like azimuth, can also be expressed in degrees. Elevation expresses



Formation of a Time Base

the angular degrees that the target is above or below the radar set.

The determination of azimuth and elevation depend upon the directional characteristics of antennas and antenna arrays. Antennas for performing these functions are discussed later in this manual.

BASIC ELEMENTS OF PULSE RADAR SYSTEMS

The six basic components in a typical pulse radar system are the timer, modulator, antenna, receiver, indicator, and transmitter, as shown in the illustration on the next page. The power supply furnishes power for all the components and may be distributed among the various components.

The Timer

The timer, or synchronizer, is the heart of all pulse radar systems. Its function is to insure that all circuits connected with the radar system operate in a definite time relationship with each other, and that the interval between pulses is of the proper length. The timer may be a separate unit by itself, or it may be included in the transmitter.

The Modulator

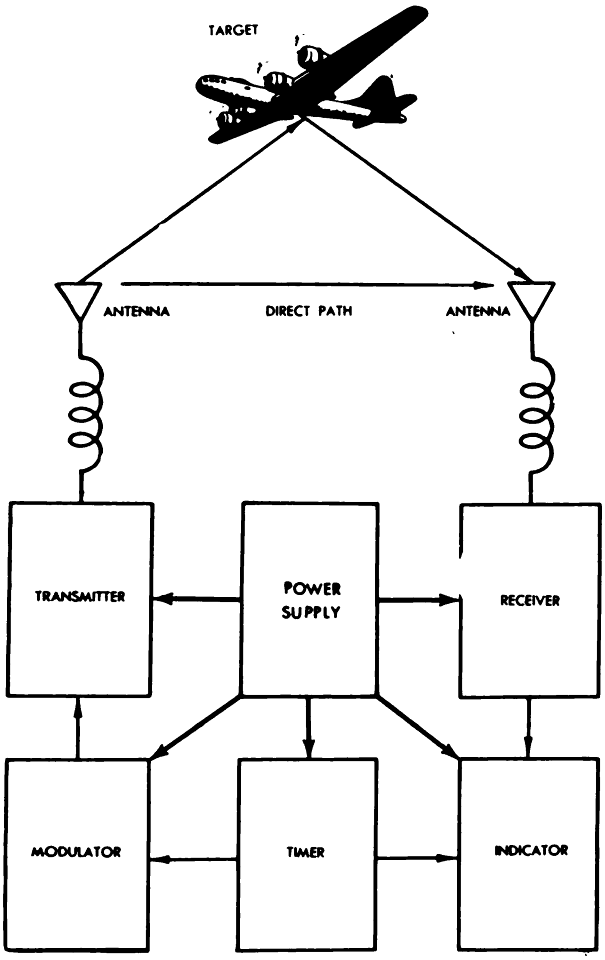
The modulator is usually a source of power for the transmitter. It is controlled by the pulse from the timer. It is sometimes called the keyer.

The Transmitter

The transmitter provides RF energy at an extremely high power in very short pulses. The carrier frequency must be extremely high to get many cycles into the short pulse.

The Antenna

The antenna is very directional in nature because it must measure the angles of elevation and the bearing of the target. To obtain this directivity at centimeter wavelengths, ordinary dipole antennas with parabolic reflectors are used. Usually, the same antenna is used for both transmitting and receiving. When one antenna is used, some kind of switching device is required for connecting the antenna to the transmitter when a pulse is being radiated and to the receiver during the interval between pulses. Since the antenna only *sees* in one direction, it is usually rotated or moved about to cover the area around the radar set. This is called *searching*. The presence of targets in the area is established by searching.



Elements of a Pulse Radar System

The Receiver

The receiver in radar equipment is usually a superheterodyne receiver. It must be capable of accepting signals in a bandwidth of one to ten megacycles. It takes the weak RF echoes from the antenna, amplifies them, and changes them to video signals which can be applied to the indicator.

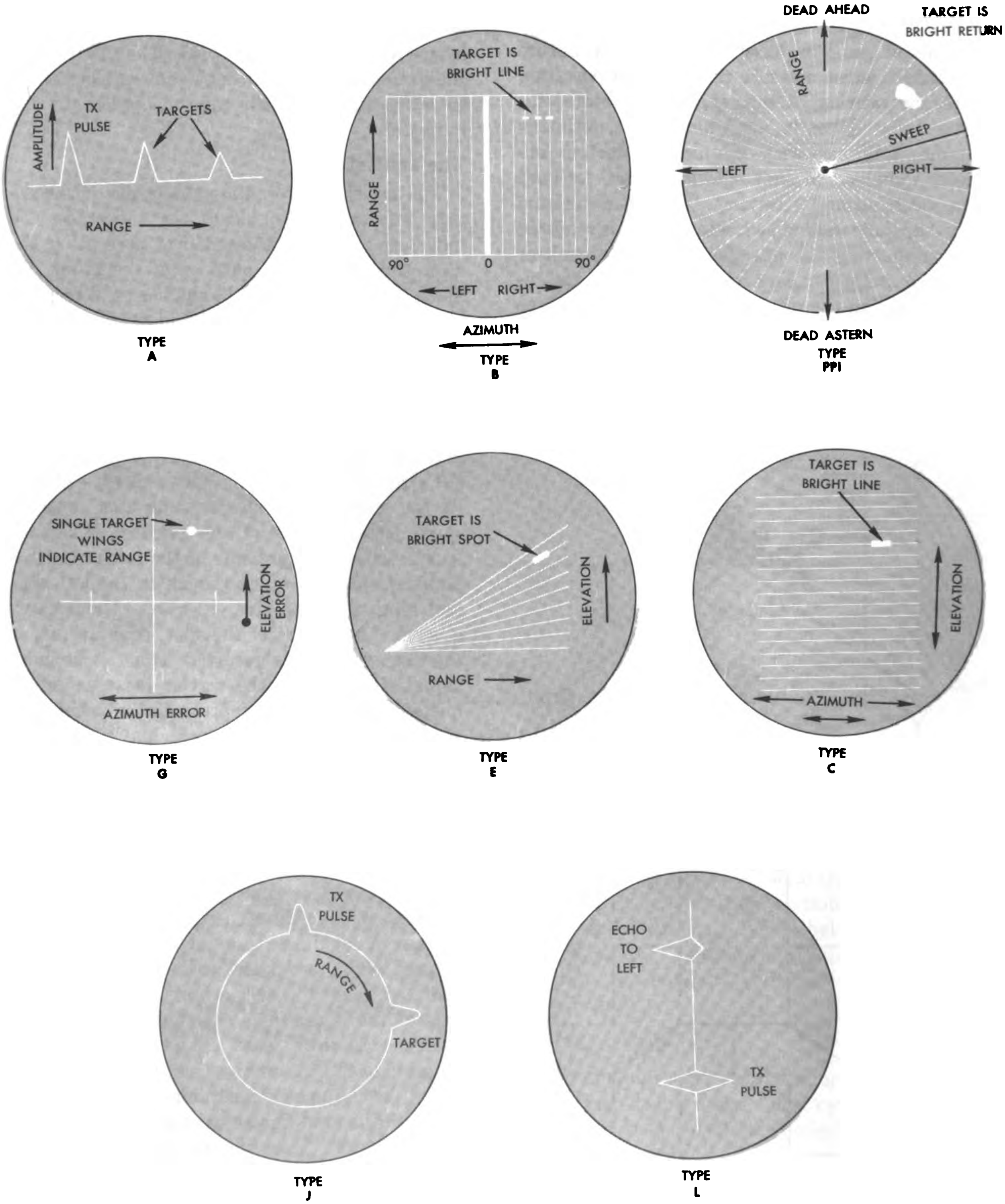
The Indicator

The indicator presents visually all the necessary information to locate the target on the indicator screen. The method of presenting the data depends on the purpose of the radar set. Since the spot scans the indicator screen to present the data, the method of presentation is often referred to as a type of *scan*.

The following common types of scan used in radar receivers are illustrated on page 328.

TYPE A SCAN. In a type A scan, the spot maintains a constant intensity, starting at the left side the instant a pulse of energy is radiated by the transmitter and traveling at a constant speed across the face of the indicator. When the spot reaches the right side of the indicator, it is blanked out, flies quickly back to the left side, and repeats the process. Receiver echoes cause a vertical deflection of the spot, proportional to the strength of the received signal. The horizontal distance on the screen between transmitted pulse and echo represents the distance to the target. Although the principal function of this scan is to determine the distance of the object, it is possible to obtain a rough approximation of direction by rotating the antenna until the maximum echo is received.

TYPE B SCAN. In a type B scan, the bearing and range of reflecting objects are presented as abscissa and ordinate, respectively. In this system, a highly directional antenna system is rotated about a vertical axis which causes the radiated beam to cover a horizontal plane. The spot on the screen has a horizontal motion which corresponds to at least a part of the angle of rotation of the antenna system. In the absence of other deflection, the scanning spot describes a bright horizontal line across the lower portion of the indicator. This line represents the transmitted pulses and is the so-called baseline of the pattern. A uniform vertical motion from bottom to top of the screen is also given to the scanning spot, each vertical line being synchronized with



TYPE A

TYPE B

TYPE PPI

TYPE G

TYPE E

TYPE C

TYPE J

TYPE L

Types of Scans

a transmitter pulse for indication of range. The vertical sweep is repeated very much more rapidly than the horizontal sweep, and the spot is maintained at low intensity. When an echo is received, the signal is impressed on the control grid of the indicator, causing a bright spot to appear on the screen. The position of this spot to the right or left of the center line of the screen indicates the azimuth of the target. The height of the spot above the baseline indicates the range of a target.

THE PLAN POSITION INDICATOR (PPI). The PPI scan is another type of scan for presenting range and bearing (direction) information. You can think of the PPI scan as a modified type of B-scan, in which the rectangular coordinates are replaced by polar coordinates. The antenna generally is rotated uniformly about the vertical axis so that searching is accomplished in a horizontal plane. The beam is usually narrow in azimuth and broad in elevation, and large numbers of pulses are transmitted for each rotation of the antenna. As each pulse is transmitted, the unintensified spot starts from the center of the indicator and moves toward the edge along a radial line. Upon reaching the edge of the indicator, the spot quickly jumps back to the center and begins another trace as soon as the next pulse is transmitted. As the antenna rotates, the path of the spot rotates around the center of the indicator screen so that the angle of the radial line on which the spot appears indicates the azimuth of the antenna beam, and the distance out from the center of the indicator indicates the range. When an echo is received, the intensity of the spot is increased considerably, and a bright spot remains at that point on the screen even after the scanning spot has passed. Thus it is possible with this type of scan to produce a map of the territory surrounding the observing station on the indicator tube. This type of scan is useful when the radar set is used as an aid to navigation.

TYPE C SCAN. In the type C scan, the echo appears as a bright spot with the azimuth angle as the horizontal coordinate and the elevation angle as the vertical coordinate. This type of scan has been used by night fighters to aid in following an enemy plane.

TYPE E SCAN. A modification of the B scan on which an echo appears as a bright spot with the range as the horizontal coordinate and the elevation as the vertical coordinate. This type of scan is used in directing planes in making a blind landing. The approaching planes must follow a certain angular line to reach the touchdown point at the left of the screen.

TYPE G SCAN. On this scan only the echo appears. It is a bright spot on which wings grow as the distance to the target is diminished. Azimuth appears as the horizontal coordinate and elevation as the vertical. This type of scan is used for pointing guns by hand. Centering the spot indicates correct aim. When the wings grow to a certain length, the range is correct for a volley from the guns.

TYPE J SCAN. A modification of the A scan in which the spot rotates in a circle near the edge of the CRT face. An echo appears as a deflection from the circle. As the distance changes, the deflection moves around the circle like the pointer of the aneroid altimeter. Its chief use is in a radar altimeter.

TYPE L SCAN. Another variation of the A scan which indicates azimuth by comparing signals from the left and right antennas. It is used in homing operations. The ratio of the signal amplitudes from the two antennas indicates the error in homing.

PULSE CHARACTERISTICS

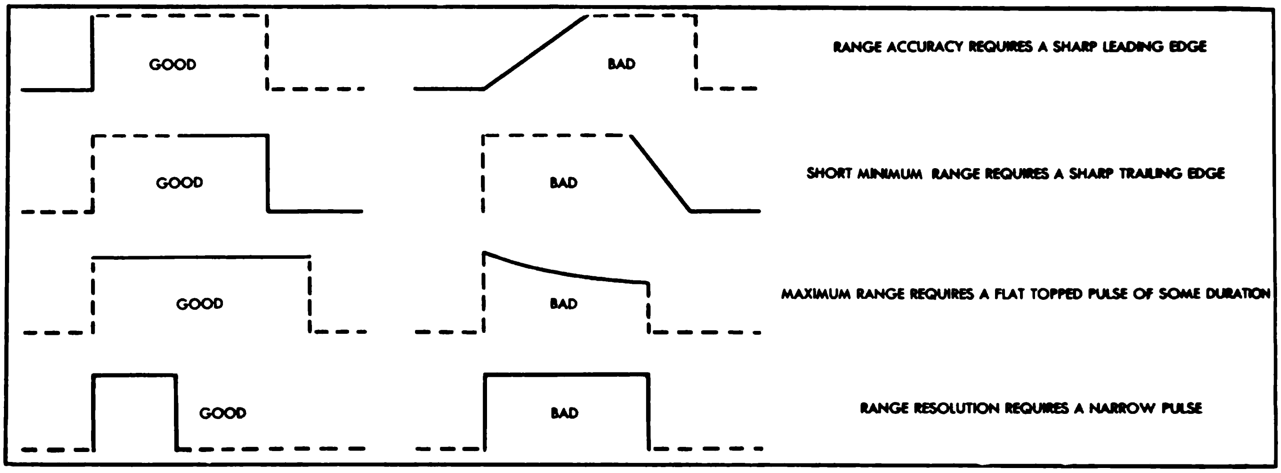
The effectiveness of a radar set on a particular mission is increased by designing it so the transmitted pulse has the proper characteristics. The most important characteristics of a radar pulse are its shape, length, repetition frequency, and power output.

Pulse Shape

The shape of the pulse determines range accuracy, minimum and maximum range, and resolution or definition of target. (See page 330.)

RANGE ACCURACY. Range accuracy depends to a large degree on the leading edge of the pulse. A pulse with a sharp leading edge provides a definite time from which to measure the start of the pulse and a better picture on the indicator scope than one produced by a pulse with a sloping leading edge.

MINIMUM RANGE. Detection of objects at the minimum range of the set requires a pulse which has a sharp trailing edge. Sloping trailing edges


GOOD
BAD
RANGE ACCURACY REQUIRES A SHARP LEADING EDGE
GOOD
BAD
SHORT MINIMUM RANGE REQUIRES A SHARP TRAILING EDGE
GOOD
BAD
MAXIMUM RANGE REQUIRES A FLAT TOPPED PULSE OF SOME DURATION
GOOD
BAD
RANGE RESOLUTION REQUIRES A NARROW PULSE

Effects of Pulse Shape

increase the width of a pulse and consequently cause inaccuracy in range since the edge of the pulse may itself cover several miles of range.

MAXIMUM RANGE. To achieve maximum range, the pulse must be flattopped and be sufficiently long in duration. A flat top in a pulse means that the energy in the pulse is constant for the duration of the pulse and all cycles reflected will have the greatest possible strength. Long pulses mean more cycles to strike the target and be reflected.

RANGE RESOLUTION. Range resolution, the ability to indicate very small differences in range where the targets are in the same direction, requires narrow pulses.

Pulse Width

Pulse width determines the minimum range of the equipment. The width of the pulse must be such that the transmitter is shut off by the time the received echo returns from the target. Therefore, the shorter the pulse, the closer to the set can echoes be received. On the other hand, long pulses provide greater maximum range since a long pulse provides more energy for the target to reflect. Long pulses also make possible a reduction in the bandwidth of the receiver. This results in less noise and an extension of range. Pulse durations used in radar equipment vary from as little as 0.25 microseconds to as much as 50 microseconds with the most common systems employing pulse durations ranging from one to ten microseconds.

Pulse Repetition Frequency

The pulse repetition frequency (PRF) largely determines the maximum range, and to some degree, the accuracy of a radar set. The actual time elapsing between the beginning of one pulse and the beginning of the next, called the pulse repetition period, is the reciprocal of the PRF. For example, if the PRF is 400 cycles per second, the pulse repetition period is 1/400 second, or 2500 microseconds. When the PRF is too high and the period between pulses too short, the echo from the farthest target may return to the receiver after the transmitter has emitted another pulse. Thus it is impossible to tell whether the observed pulse is the echo of the pulse just transmitted or the echo of the preceding pulse. Such a condition is referred to as *range ambiguity.*

Although the pulse repetition rate must be kept low enough to make possible the required maximum range, it must also be kept high enough to avoid some of the pitfalls a single pulse might encounter. If a single pulse were sent out by a transmitter, atmospheric conditions might attenuate it, the target might not reflect it properly, or moving parts, such as a propeller, might throw it out of phase or change its shape. If it did return, you might not see it at all. Thus, information derived from a single pulse would be highly unreliable. By sending many pulses, you will get many good reflections from the target. The equipment will integrate or sum up the good points of all the pulses and present you with a good clear reliable picture. Equipment is therefore designed in such a way that many pulses, ten or more, are received from a single target. In this way, the effects of fading are somewhat reduced. At short ranges, such as those used in gun-laying sets, the repetition frequency is increased in order to insure accurate short-range measurements. Many echoes are therefore received from each target, and the integrating effect is increased. This makes possible the more accurate range and bearing determination necessary in gun-laying equipment.

Another factor affecting the PRF is the rate of angular motion of the antenna in searching extensive regions. If the antenna is moved through too large an angle between pulses, not only will the number of pulses-per-target be too low, but there may even be areas in which targets may exist without being detected. In this connection, still another factor to consider is the sharpness of the antenna beam. A sharp beam must be pulsed more often than a wide beam to avoid skipping over targets.

Tactical employment of a radar set determines, to a large degree, the PRF to be used. Long-range search sets, covering distances to 150 miles, require a pulse rate slow enough to allow echoes from targets at the maximum range to return to the receiver before the transmitter is again pulsed. Higher pulse rates are used in aircraft interception sets where the maximum range is around 10 miles.

Power Requirements in Pulse System

The maximum range which a radar set can attain depends largely on the power output of the set. Enough power must be radiated so that at the maximum range, the received echo signal will have a power level at least equal to the electronic noise level at the receiver.

Radar systems have been developed to the degree where they transmit the largest peak power ever radiated by any type radio transmitting equipment. This peak power is sometimes as high as 5 megawatts, and seldom less than 20 kilowatts.

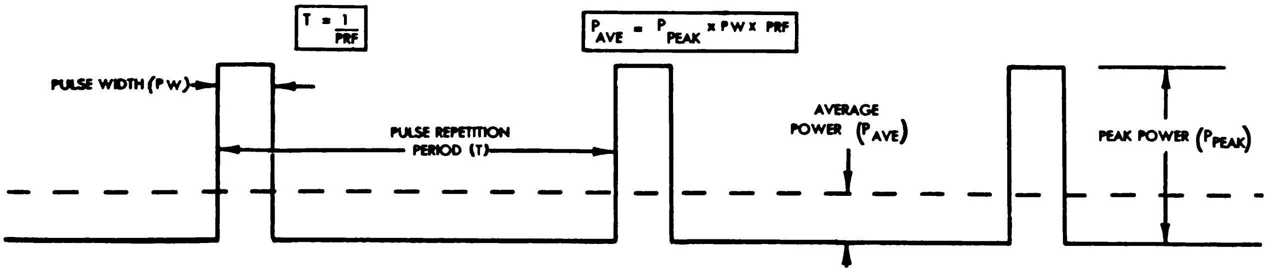
In considering power requirements, you must distinguish between two kinds of power output — *peak power*, which is the power during a pulse, and *average power*, which is the average power over the pulse repetition period. While the peak power is very large, the average power may be small, because of the great difference between pulse width and pulse repetition.

The ratio of the pulse length to the pulse repetition period is known as the duty cycle. To convert peak power into average power, multiply the peak power by the duty cycle. These terms are defined in the illustration below. For the pulse waves shown, the relation between the peak power and the average power is expressed by the equation,

$$P_{ave} = P_{peak} \times PW \times PRF$$

where PW is the pulse width in seconds.

The amount of power that a transmitter must radiate depends first on the minimum amount of power to which the receiver will respond, and second, on the elements which govern attenuation of power to and from the target. Such


$T = \frac{1}{PRF}$
$P_{AVE} = P_{PEAK} \times PW \times PRF$
PULSE WIDTH (PW)
PULSE REPETITION PERIOD (T)
AVERAGE POWER (P_{AVE})
PEAK POWER (P_{PEAK})

Pulse Characteristics

↑ **Increase**

→ **No Change**

↓ **Decrease**

PRF	PULSE WIDTH	PEAK POWER	AVERAGE POWER	MAXIMUM RANGE
(↑)	→	→	↑	↓
→	→	(↑)	↑	↑
→	(↑)	→	↑	↑
→	(↓)	(↑)	→	→

Circle denotes the elements (characteristics) which are varied. For example, if the PRF is increased, then the average power must be increased to maintain the same peak power. This results in less maximum range (less listening time).

Effect of Varying Pulse Characteristics

elements are moisture in the atmosphere, clouds, and weatherproof covering over the antennas. The effect of varying the pulse characteristics of a radar set is summarized in the table shown.

HOW RADAR WAS DEVELOPED

In the first part of this chapter we've introduced you to the technical background of radar: we've explained its kinship with radio and the physical laws on which it is based; we've defined some of the terms which the operator uses frequently and told you the principal units which make up a radar set. But this survey might leave you with the impression that radar was hatched overnight into a full grown electronic chicken. Because it wasn't, we're going to pull some sheets off the calendar and see how the marvelous gadget called radar grew to its present stature.

Since radar is the direct descendant of radio, we're sure you won't be surprised when we tell you that two types of radio equipment developed during the two decades before World War II furnished the basic ideas for radar. The first was a compact type of direction-finding equipment which uses directional antennas. The other — you guessed it — was television, which uses the cathode-ray tube.

Let's pull a lot of sheets off our calendar now, back to 1922. For in that year Dr. Albert M. Taylor of the Naval Research Laboratory first observed "radio echoes." That is, he observed that a ship passing between a radio transmitter and a radio receiver reflected some of the waves back toward the transmitter. Between 1922 and 1930, further tests proved that surface vessels hidden by smoke, fog, or darkness could be detected by radio waves. Because the discovery had great military value, further experiments were strictly hush-hush. During this same period, 1922-1930, Doctors Breit and Tuve of the Carnegie Institute published reports on the reflection of pulse transmission from electrified layers of the upper atmosphere. You can imagine how important that study was. From their discovery followed the application of the principle of reflected pulse transmision to the detection of aircraft.

We didn't have far to go then. By 1930, the National Bureau of Standards Radio Division had developed a radio set which could detect an aircraft in flight. And finally, in 1934, further research produced a single set which could determine range, the primary accomplishment of radar. In the same year the U. S. Navy developed radar sets for ship-to-shore installations.

Meanwhile, our cousins across the sea, the English, hadn't exactly been asleep either. Robert Alexander Watson-Watt had developed what was essentially a radar set for some experiments he was conducting with a layer of the upper atmosphere called the ionosphere (which begins about 25 miles up). During these experiments, radio echoes were picked up at ranges too close to be attributed to the ionosphere. Finding that these echoes were caused by aircraft and realizing the military value of his discovery, Watson-Watt told the Air Ministry what he had found. Soon after, in 1935, radar was first demonstrated. With slightly modified equipment placed in a truck, the scientist was able to detect an approaching airplane and follow it in its flight.

Radar as a Defensive Weapon

Unfortunately, it wasn't long before the English had to put this new equipment to practical use. As Hitler moved against Austria and Czechoslovakia, it became obvious that general war was almost inevitable, and the Air Ministry decided to adopt this new defense against bombing. During the years before the war, the production of early-warning sets and their installation along the English coast took priority over most other defense measures. With feverish haste, CH (Chain Home) stations were erected in a net which faced Germany. When war was declared, an air defense system using these stations was set up. By the time the Battle of Britain began, England was ready at least in this respect.

Although much progress had been made in ground radar, not much encouragement was given to the development of airborne radar equipment until April 1940. The RAF Coastal Command, using airborne equipment, sighted the missing German prison ship, *Altmark*, in a Norwegian fjord. Once the Altmark incident had established its value, however, airborne radar was quickly adopted as nightfighter equipment.

The requirements of airborne equipment led next to microwave radar (microwaves are short electromagnetic waves), perhaps the greatest advancement radar made during the war. This development made airborne equipment even better, as microwave radar was very accurate. It meant, too, that radar could now be employed for offensive bombing and gun-laying, greatly in demand after the Allies had taken the offensive.

In the last section, we suggested the value of radar as a defensive weapon. Let's carry that a little farther before turning to its other important characteristic, its usefulness as an offensive weapon.

One thing that may surprise you about radar — it's more durable that you'd think. During the Battle of Britain, stations in southeast England were bombed and their reception interfered with, but they still did their job well. They could, for example, warn of enemy raids hours ahead of time because they could detect German aircraft, 500 or more of them, flying in circles over Belgium or northern France as the Luftwaffe grouped for its raids.

But the radar installations did much more than just let the cities know they should get ready for a raid. By using the information gathered by radar, the fighter command of the RAF was able to cripple the Luftwaffe by having fighters at the right place at the right time. Through early-warning detection and location of enemy aircraft by radar installations, intercepting fighters were placed directly in the path of the approaching bombers, no matter where they attacked from. One Sunday, 15 September 1940, the Luftwaffe lost over 100 aircraft to RAF fighters guided by radar; from that Sunday on, a black Sunday for the Germans, the daylight raids on Britain declined.

Radar as an Offensive Weapon

You can see, then, that radar played a mighty important part in the air battles over Britian. Yet it was not only in defending England from the bombers of Germany that radar played an important part; it also proved to have offensive value. The Germans we know were aware of that importance, for crude radar equipment was found aboard the scuttled *Graf Spee*. Elsewhere on the sea radar showed its worth. At Taranto, the British navy dealt the Italian fleet a severe blow; and during landing operations in Africa, the Americans damaged the *Jean Bart*. Radar, though, fully matured with the development of the Combat Information Center as the operations room of a warship. These CIC's used radar as their main source of information. In the Pacific, we won several major victories by effective use of this tactical weapon.

At Pearl Harbor, Japanese airplanes were detected by radar when they were more than half an hour away. To our sorrow, they were confused with some American planes that were expected, and no one paid any attention to the warning. Later, radar helped us against the Japs. At Midway, the use of ship and ground radar enabled airplanes to disperse and defeat a far superior Japanese force.

Halfway around the world, and later, the Anzio beachhead was relieved from German night bombing by the employment of a new microwave gunlaying radar after all other radar equipment had been neutralized by jamming. The use of H2X bombing equipment on Eighth Air Force Flying Fortresses made high-precision bombing of German targets possible even in bad weather.

So you can understand why radar, successful as it proved, was completely accepted during the final phase of the war. We used it so often and in so many ways that its achievements were just routine. No longer did we just point to isolated incidents of jobs done well; we knew we had a weapon we could depend on.

The last few months of the war continued the radar success story. By V-E day ground radar control centers and navigational aids such as SHORAN (short-range navigation) were common. Tactical air commands used radar equipment cooperatively. To this equipment were added mortar detection equipment and moving target indicators for the proposed invasion of Japan. But that invasion never came off. Bombing by B-29's, heavily armed with the latest in airborne electronic aids for bombing, RCM (radio countermeasures) equipment, and navigational aids, such as LORAN (long-range navigation), brought V-J day without costly assault.

PRESENT USES OF RADAR

Back in the preceding section, we touched on radar's importance in the defense of Britain. Suppose we examine now in more detail just how radar is useful in defense and offense.

Part of its usefulness relates to the speed of modern airplanes. Fast as they are, if they are hostile we have to know their height, bearing, and range long before we can see them or hear them. Similarly, fog, clouds, or smoke, either during day or night, must not interfere with detecting enemy airplanes. So the fact that any enemy today would have extremely fast aircraft means that your detector must be fast too. We've already seen that the waves which radar uses travel at the speed of light, and so radar is fast enough.

Height Finding

As soon as we found out we could blanket an area with radio energy so that no object could pass through the area without our knowing it, we devised methods for indicating the distance and direction of the object. These we've discussed in earlier sections. However, by applying known physical laws, scientists went even further. They developed *height-finding* equipment accurate enough to determine the elevation of distant aircraft to within 500 feet. The basic principle used was to place two sets of antennas at different heights above the ground and compare the strength of their echoes. This three-dimensional location system, with techniques for estimating the total number of aircraft in flight, provides necessary information for establishing defense plans.

Identification

There is one other very important piece of information needed, besides the height, range, and direction of aircraft. You have to know whether the approaching aircraft are hostile or friendly. We solved this problem by making the electrical responses or echoes received from our own planes different from those received from enemy airplanes. Devices have been developed and adopted which identify friendly aircraft by automatically coded replies which accompany the reflected echoes. Known as IFF (Identification, Friend, or Foe), they are pretty remarkable inventions.

Fighter Equipment

What ground radar can do in air defense is naturally very important. Once hostile aircraft have been detected, however, we need to send interceptors up to attack them. These interceptors need miniature radar stations to aid them in finding and firing on the enemy, even in clouds or at night. Since this equipment must fit into a small aircraft, it must be small and light, and its antennas must not project or they will reduce air speed. Another consideration is that the pilot is going to have his hands full flying the airplane and fighting and can't take time for complicated adjustments and calculations. So you see the airborne set in a fighter plane must not only be small but simple to operate as well.

In our early fighter radar, pilots temporarily lost night vision after looking at the screen for a long period. Eventually, however, even this problem was whipped — by reflecting the indications on the large end of the tube onto the airplane windshield at such an angle and focused so that as the pilot looked ahead he saw a ghostly green flow representing the radar echo. As the enemy approached, the glowing spot increased, *growing wings* as it came within firing range. Finally, the pilot would see the aircraft as a dark shape in the night sky, just where the green image had appeared.

Navigational Aids and Bombing

Once we found that airborne radar would work effectively, we soon thought of using radar as a navigational aid, that is, using the radar set in the aircraft to receive echoes from beacons. That was just the beginning. Soon lightweight portable beacons enabled navigators of troop carrier planes to find the unloading area in any weather. These portable radar sets, Eurekas, were used in the African landings.

Nor was this the end of what could be done with radar. When we began to bomb German targets, radar type navigational equipment became very important. The GEE system was prepared by the British in 1942 and, along with other pathfinder equipment, was used successfully in the great air offensive that began with the first 1,000-plane raids.

Offensive Ground Action

Radar solved other problems, too. The difficulty of getting adequate aircraft warnings coverage for offensive ground actions and task forces was overcome in 1942 by Radio Set SCR-602-T1, the first of the lightweight, portable radar sets which became standard equipment for forward areas. Take also the matter of gunlaying. While microwave gunlaying equipment was developed during the first years of the war, we didn't get high-accuracy and low-altitude coverage until high-powered microwave aircraft warning equipment was perfected. So much information could be obtained by using this equipment that elaborate operations rooms, called control centers, were placed near the sets to control fighter sweeps and other operations. Ground radar could indeed be employed for purely offensive operations.

So you can see that radar contributed greatly to our winning the last war. Because it partly eliminated the hazards of flying in miserable weather, greatly increased the target hits of longe-range bombing missions, and increased the carrying of troops by air and the air transport of freight, radar equipment (like paratrooper-operated beacons, LORAN, SHORAN, AND GCA) was produced as speedily as possible. Radar sold itself during the last war and continues to sell itself everyday by its many new uses.

SUMMARY

We've certainly talked a lot about the different kinds of jobs radar can do. We believe you'll agree that most of them fall into one of three big groups: search, fire control, or navigational aids.

Let's review briefly *search*, whose primary function is to find targets and provide continuous information about their location. With this information, we can estimate the situation and decide what counteraction to take. Successful search equipment must operate over a long range, provide complete coverage of an area, and make its information easily visible. To do this, you recall, a search radar set uses pictorial oscilloscope scans such as PPI or type A.

Closely related to search is *fire control*. Fire-control equipment provides accurate locating information to be used in actually firing on specified targets. Certainly, this equipment must be very accurate and must include some mechanical or electrical means to convert and transmit the information into fire-control computing instruments, such as antiaircraft artillery gun directors. Range, coverage, and pictorial presentation are not so important in this equipment since the operator doesn't need them — the computer is the middleman on this deal. A desirable, though not always necessary, feature of fire control equipment is automatic tracking, the ability to keep the gun automatically directed on the target.

The third important use of radar is to help ships and planes *navigate* better. Nearly all radar equipment can do this work. Again, what's needed is long range, complete coverage, and a good picture. Accuracy may be essential, although how accurate depends on the specific type of aid wanted.

Radar obviously is much more than just a weapon of war; it also has a bright peacetime future. One use is to help commercial ships and aircraft operate more safely. Besides that, the growing up of radar has stimulated the entire science of electronics. The vacuum tube circuits designed for radar are capable of hundreds of new peacetime uses.

That pretty well takes care of our introduction to radar. As you study radar, you should remember that in spite of its newness and the air of mystery which has surrounded it, radar is primarily a radio device which uses the principle of wave reflection to locate targets. The main parts of a radar set are the transmitter, the receiver, the antenna, the indicator, the timer, and the power supply. The general uses of radar are in search, fire control, and navigation.

Chapter 34

Microwave Oscillators

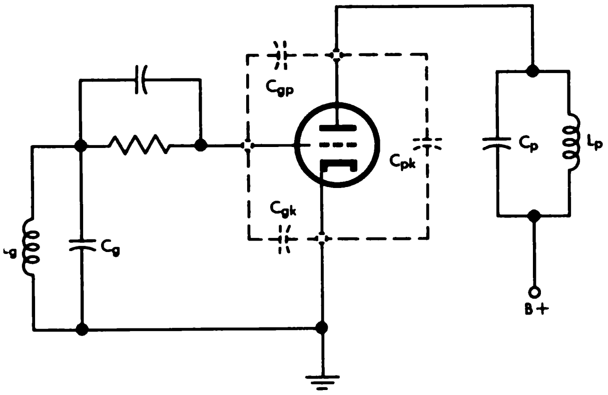
The conventional oscillators studied in chapter 27 have limitations when used at radar frequencies. These limitations are the result of many factors, some of which are due to the physical structure of the tube and others, to the external circuits.

Effect of Physical Structure

The physical structure of a tube limits the operating frequency of a conventional oscillator because of the interelectrode capacitances present in the tube, the inductances present in the leads, the transit time effects, and dielectric losses.

INTERELECTRODE CAPACITANCES. At ordinary radio frequencies, the interelectrode capacitances of a vacuum tube introduce reactances which are large; hence, they do not have any serious effects. As the frequency increases, however, the reactances presented by the interelectrode capacitances become less; therefore, more current is drawn in charging these capacitances. For example, an rf oscillator which has a plate-to-grid of capacitance 10$\mu\mu$f, presents a reactance of 159-k ohms at 100 kilocycles and of only 159 ohms at 100 megacycles. If the rf voltage between plate and grid is 500 volts, the current drawn in charging this plate-to-grid capacitance is 3.15 amperes at 100 megacycles and only 3.15 ma at 100 kilocycles. These high charging currents at the higher frequencies dissipate more power than is being supplied as power current by the tube.

Another way in which interelectrode capacitances limit the operating frequency of conventional oscillators may be seen from a study of the illustration on the next page. The interelectrode capacitances are shown in the dotted lines. Note that the plate-to-cathode capacitance is in parallel with the series combination of the plate-to-grid capacitance and the grid-to-cathode capacitance. All these capacitances effectively form a part of the capacitance of the tuned circuit.



A TPTG Oscillator with Interelectrode Capacitance

Thus, it is seen that, no matter how small the capacitance is made, the highest frequency generated is limited by the interelectrode capacitances. In addition, the interelectrode capacitances vary with the applied voltage and the loading of the oscillator, causing frequency instability. This is especially true when C is very small and the interelectrode capacitances form a large part of the total capacitance.

INDUCTANCES IN THE LEADS. Another factor which limits the frequency of a conventional oscillator is the inductance present in the lead to each element. These inductances, which have negligible reactance at low radio-frequencies, begin to have considerable reactance at higher frequencies. Since the inductance of the cathode lead is common to both grid and plate circuits, its reactance causes degenerative feedback which reduces the amplitude of oscillations at the higher frequencies. In addition, as the inductance is decreased in order to generate a higher frequency, the lead inductances represent a greater portion of the inductance and set a limit on the highest frequency that may be generated.

TRANSIT TIME. The physical structure of a tube also sets a limitation on the frequencies that may be generated by a conventional oscillator because of transit time. Transit time is the time required for electrons to travel from the cathode to the plate of a tube. In low-frequency operation, this time is disregarded because it is negligible when compared with the time of one cycle. Thus, grid voltage and plate current are considered to be in phase and plate voltage to be 180° out of phase with them. At higher frequencies, however, this transit time becomes appreciable when compared with a wavelength. For example, if the transit time of a conventional tube is on the order of 0.001 microsecond, this time represents a half-cycle at 500 megacycles but only 1/2000 of a cycle at 500 kilocycles. Hence the lag of plate current behind grid voltage becomes an appreciable portion of a wavelength at the higher frequencies. This lag in plate current causes the plate voltage and grid voltage phase relationship no longer to be 180°. As a result, oscillations cannot be maintained in conventional oscillators at those frequencies where transit time is excessive. It has been found experimentally that at those frequencies where transit time is greater than one tenth of a cycle, the efficiency of a conventional oscillator drops considerably. When transit time approaches a quarter of a cycle, the oscillator does not oscillate at all.

DIELECTRIC LOSSES. Other losses due to the construction of the tube which become especially objectionable at high frequencies are dielectric losses between the tube base and the envelope. These losses reduce the output and efficiency of conventional oscillators when used at high frequencies.

Effect of External Circuits

As frequency increases, conventional oscillators also have limitations which are due to the external circuits of the oscillator. These limitations are caused for the most part by two types of losses: skin effect and radiation. An increase in frequency causes a considerable increase in the resistance in a vacuum-tube circuit because current flows in a thin layer on the surface of a conductor. The depth of the layer becomes thinner as the frequency is increased. Hence, as frequency increases, the I^2R losses due to skin effect become greater.

Radiation, which is the other cause of power loss in a conventional oscillator at high frequencies, is due to the incomplete cancellation of electromagnetic fields in the region surrounding the circuit. At low frequency a spacing between conductors which represents only a small fraction of half-wavelength causes complete cancellation of fields in all directions. As the frequency increases, however, the identical spacing represents a larger fraction of a half-wavelength, causing less cancellation. At the frequency at which the spacing represents a half-wavelength, the fields add in the direction of the plane carrying the two conductors, causing radiation as though from an antenna.

ULTRAHIGH-FREQUENCY TUBES

The limitations due to the physical structure of tubes may be overcome by constructing tubes of special types. Low interelectrode capacitance may be attained by means of small, carefully designed electrodes. Although the reduction of interelectrode capacitances is made possible by an increase in the spacing between electrodes, it also has the adverse effect of increasing transit time. Hence, ultrahigh-frequency (UHF) tubes tend to have small electrodes, a condition which causes them to have low power capabilities.

Reducing the physical size of the tube and making the leads shorter and of larger diameter can reduce the lead inductances. The leads in most (UHF) tubes are brought out through the tube envelope, without a conventional tube base being used. This also reduces dielectric losses if the leads are brought through the glass at points where there are voltage nodes.

Transit time may be reduced by closer spacing of electrodes and higher plate voltages. Bringing electrodes closer together, however, increases the interelectrode capacitance of the tube. This method is, therefore, used only if electrodes are made smaller at the same time. In most cases, higher plate voltages are applied to reduce the transit time.

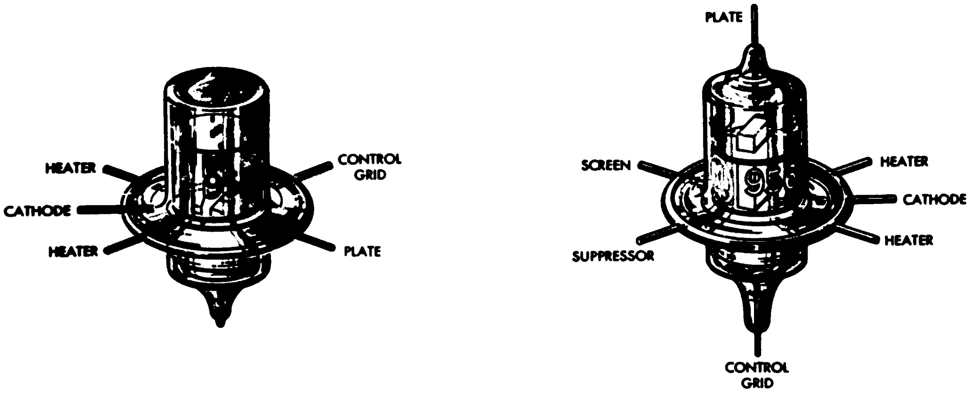
Acorn Tubes

Two acorn tubes, one of which is a triode and the other a pentode, are shown in the illustration below. The internal arrangement of these tubes is similar to that of a conventional tube, but they have small spacing between electrodes to reduce transit time, small electrodes to reduce interelectrode capacitance, and short thick leads to reduce lead inductance. Note that the leads are brought directly through the glass and that no conventional tube base is used. Tubes of this type are designed for use as oscillators or amplifiers at frequencies up to about 300 megacycles. Frequencies up to 600 megacycles may be attained, but efficiency is very low at this frequency. These tubes are called acorn tubes because of their appearance and small size.

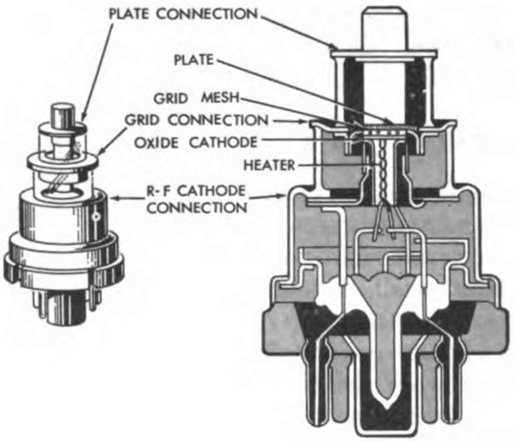
Lighthouse Tube

Another single-tube oscillator used at frequencies as high as 3600 megacycles is the lighthouse-tube oscillator shown in the illustration on page 340. Because the peak power output is limited, this type of oscillator is used as a local oscillator or as a transmitter in low-power radar equipment.

The lighthouse tube is usually mounted within a tuning assembly composed of three coaxial cylindrical conductors, shown in the illustration on page 341 of a cross section of a lighthouse tube and tuner used as a transmitter.



Acorn Tubes: A. Triode B. Pentode



A Typical Lighthouse Tube

The inner conductor makes contact with the plate, the next with the grid, and the outer one with the shell of the tube, which is connected to the cathode both externally and through interelectrode capacitance. The space between cathode and grid conductors forms a coaxial cathode line, short-circuited by an adjustable plunger. The plunger does not actually touch the grid conductor, but the capacitance between the two provides an rf short circuit. The DC grid-to-cathode path consists of R_g and C_g and forms the grid-leak bias circuit.

The grid and plate conductors form the plate line, which is open-circuited at the end away from the tube. Plate voltage is applied to the plate through the plate tuning rod at the point where the plate line is short-circuited. The quarter-wave section of line behind this short acts as an rf choke because it presents a high impedance and keeps rf out of the power supply.

The short cathode line acts as an inductance, since it is slightly less than ¾ of a wavelength. The open-circuited plate line also acts as an inductance, since it is slightly less than a wavelength. These dimensions are chosen for a frequency of approximately 3000 megacycles, but the shorted cathode line would act as an inductance just as effectively if it were slightly less than a quarter-wavelength, or any odd multiple of a quarter-wavelength, at any frequency; and the open-circuited plate line would act as an inductance just as effectively if it were slightly less than a half-wavelength. You will see why this is so when you study chapter 35. Since the lines are connected directly to the electrodes, there are no lead inductances to be considered.

To summarize, for oscillations to occur in a lighthouse-tube oscillator, the plate line must be tuned inductively and the cathode line must be tuned capacitively. Thus, the plate circuit must be tuned to a higher frequency than the cathode circuit, and the oscillations which occur are at a frequency somewhere between the two. The tuning of the cathode line controls principally the amount of feedback voltage; the tuning of the plate line controls principally the frequency of oscillation. Because of the interaction between these tuning controls, it is often necessary to readjust each line to obtain maximum output at the desired frequency.

THE MAGNETRON OSCILLATOR AS A TRANSMITTER

Practically all radar transmitters which operate above 600 mc use magnetrons for transmitting tubes. The main advantage of the magnetron is that transit time, which is a determining factor in the upper limit of frequency of operation in conventional tubes, is not a disturbing influence. There are some special type triodes employed in unusual transit time oscillator circuits that will oscillate in the frequency range useful for radar, but unlike the magnetron — their power output is too small for them to be useful in most cases. In addition, there are also the velocity-modulated tubes, such as the klystron and the Shepherd-Pierce tube.

These operate in the desired frequency range with better power output than the triodes just mentioned but they are not readily adaptable for pulsed operation.

Physical Description of a Magnetron

The magnetron is a diode in which the magnetic field between the cathode and the plate is perpendicular to the electric field. The tuned circuits included in the tube are in the form of cylindrical cavities. As shown in the illustration of a typical magnetron on page 342, the cathode and filament structure is at the center of the tube and is supported by filament leads. These filament leads are large and stiff enough to keep the cathode and filament structure fixed in position under ordinary circumstances.

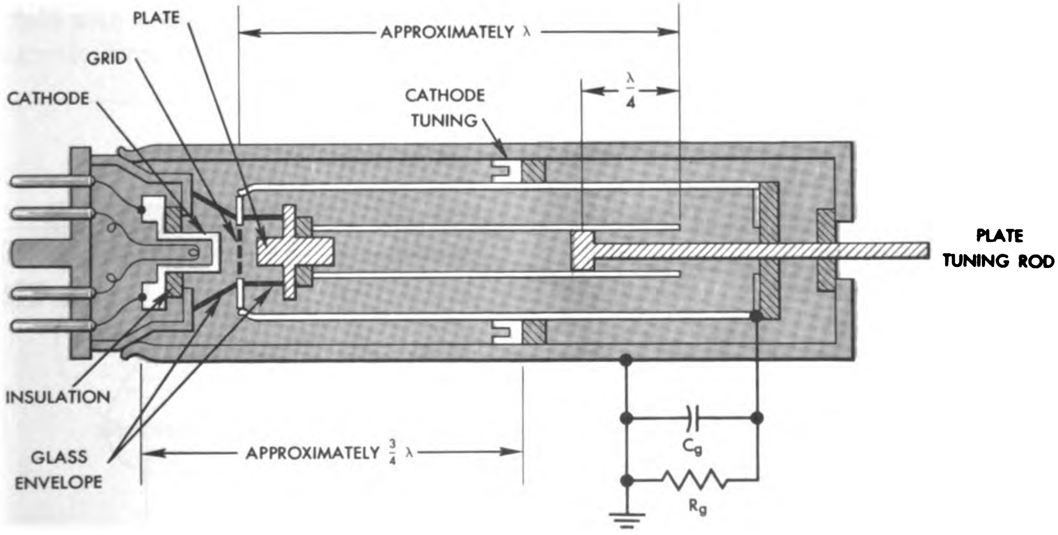
The plate structure, as shown in the cutaway view on page 342, is a solid block of copper in which the resonant cavities are cylindrical holes. A narrow slot opens each cavity into the central portion of the tube and divides the inner plate structure into as many segments as there are cavities. Alternate segments are strapped together to put the cavities in parallel so far as the output is concerned. This makes it possible to take the output from a pickup loop placed inside any one of the cavities. Since the outer conductor of the output coaxial line is connected to the shell of the magnetron, which is a part of the plate structure, there cannot be a high, positive DC voltage applied to the plate. Instead, a negative voltage is applied to the cathode to accomplish conduction. The magnets, which are made of a special alloy called Alnico, are unusually strong.

Motion of Electrons in Electric and Magnetic Fields

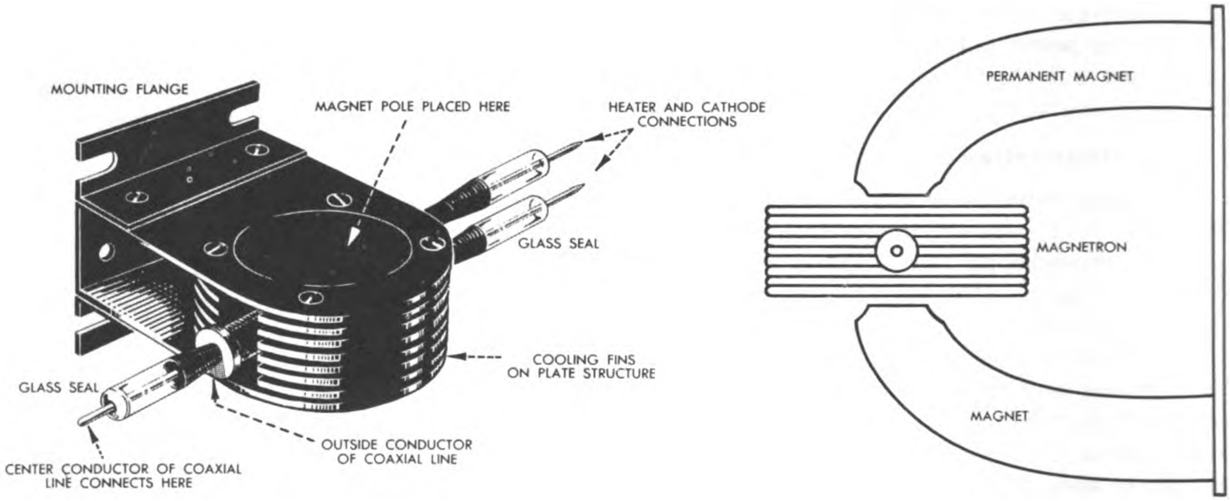
The theory of operation of the magnetron is based on motion of electrons in combined electric and magnetic fields. The following laws govern this motion.

IN ELECTRIC FIELD ALONE. The law governing the motion of an electron in an electric field states that the force exerted by an electric field on an electron is proportional to the strength of the field, and the direction of the force is opposite to that of the field. In other words, electrons tend to move from low potential toward high potential.

IN MAGNETIC FIELD ALONE. The law of motion of an electron in a magnetic field states that the force exerted on an electron moving in a magnetic field is at right angles to both the field and the path of the electron. The direction of the force is such that the electron travels clockwise when viewed in the direction of the magnetic field as shown in the diagram on page 343. Here the direction of the magnetic field is into the page. A stationary electron in a magnetic field is not affected by the field.



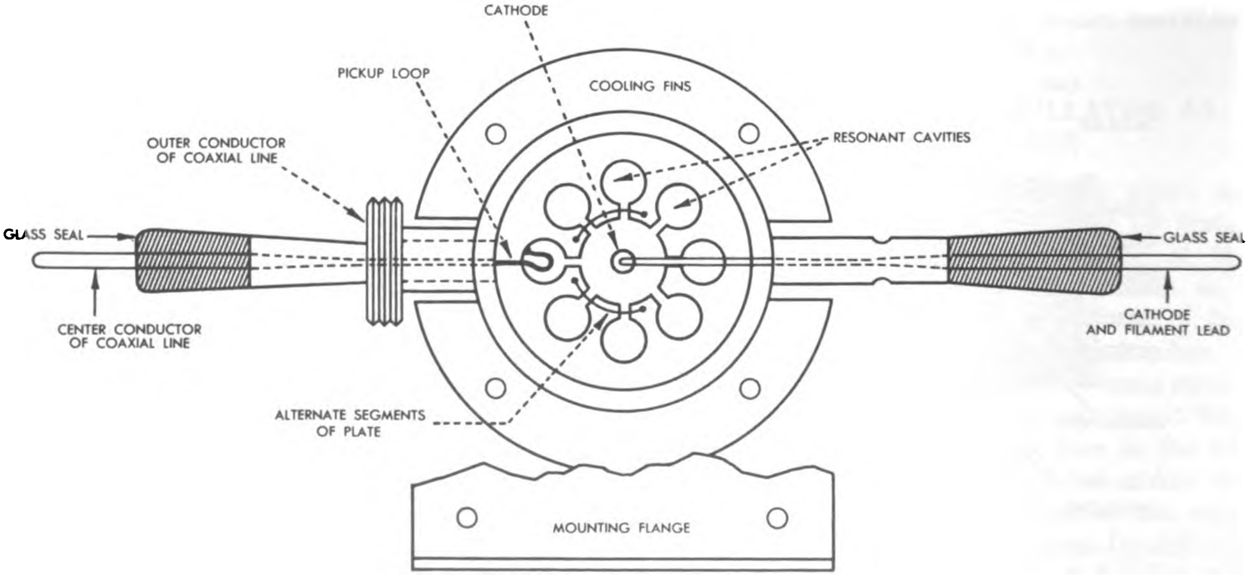
A Lighthouse Tube and Tuning Assembly


MOUNTING FLANGE
MAGNET POLE PLACED HERE
HEATER AND CATHODE CONNECTIONS
GLASS SEAL
COOLING FINS ON PLATE STRUCTURE
GLASS SEAL
OUTSIDE CONDUCTOR OF COAXIAL LINE
CENTER CONDUCTOR OF COAXIAL LINE CONNECTS HERE
PERMANENT MAGNET
MAGNETRON
MAGNET

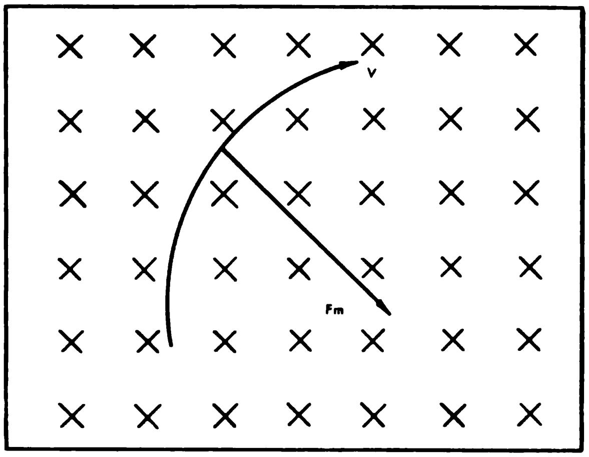
A Typical Magnetron

In mutually perpendicular electric and magnetic fields. The path that an electron takes when traveling perpendicular to a magnetic field depends upon the initial velocity of the electron. Generally, however, the paths of electrons in a plane perpendicular to the magnetic fields are cycloids (curves generated by points on a radius of a wheel rolling on a smooth plane). Different paths are generated by different points along the radius of the wheel. A point on the circumference generates a typical cycloid as shown by point C in the illustration at the bottom of the next page. Points lying beyond the radius (d) generate cycloids also, but these are closed loops.

All paths generated by the points on the radius of the rolling wheel may also be generated by electrons traveling at different velocities. You can understand this if you assume that the uniform magnetic field is directed inward perpendicular to this page and that the electric field is directed downward from the top to the bottom of the page. If an electron is projected horizontally from left to right, there will be


CATHODE
PICKUP LOOP
COOLING FINS
OUTER CONDUCTOR OF COAXIAL LINE
RESONANT CAVITIES
GLASS SEAL
GLASS SEAL
CENTER CONDUCTOR OF COAXIAL LINE
CATHODE AND FILAMENT LEAD
ALTERNATE SEGMENTS OF PLATE
MOUNTING FLANGE

Cutaway View of Magnetron



Path of an Electron Moving in a Magnetic Field

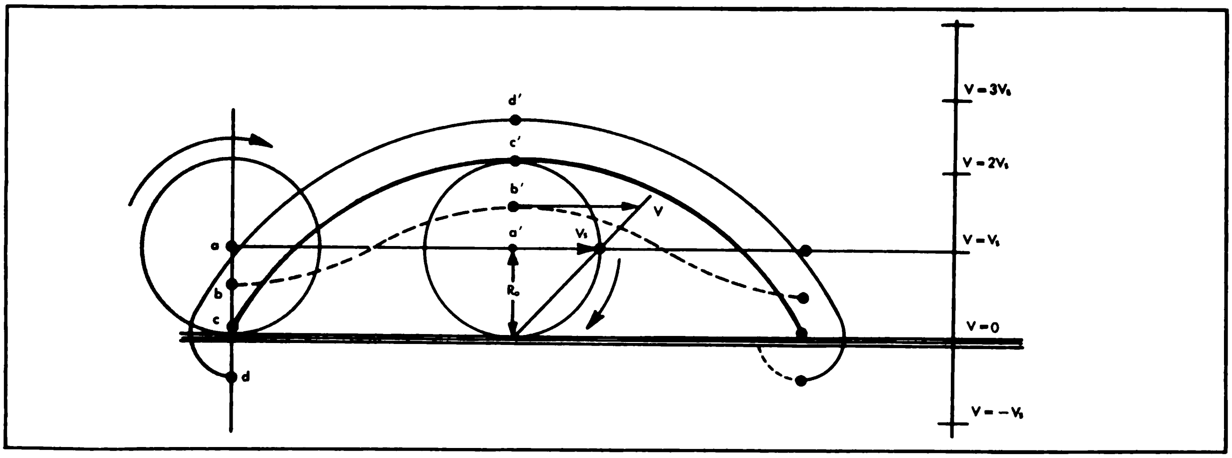
two forces acting on it — a magnetic force downward and an electric force upward. If these forces are equal, the path of the electron E will be a straight line as shown by electron *a* traveling at velocity V_s in the illustration on the next page.

If the forces are unequal, the electron path will be a cycloid as shown in the same illustration. All electrons traveling with velocities between zero and V_s take a path similar to that of electron *b*. The qualitative explanation of how this path is generated is the following: Because the initial velocity of the electron is less than V_s, there is a smaller magnetic force acting on it. Hence, the electron curves upward at first. But as it moves, it gains kinetic energy from the electric field with a resultant increase in velocity. The magnetic force thus increases and after a time the electron path is bent downward where the electron begins to lose velocity, and finally returns to its original state.

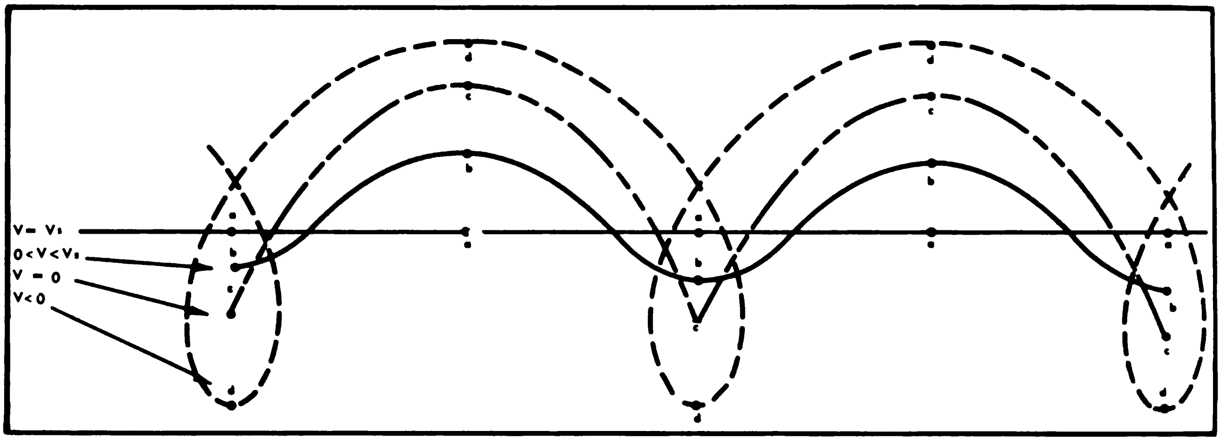
Curve *c* above shows the path which an electron takes when it starts with a zero velocity. At the start, the magnetic field exerts no force on the electron as no magnetic lines of force are being cut. The electric field, however, moves the electron directly upward to a point where the magnetic field acts on it and deflects the electron to the right. As the electron moves upward, it gains velocity, the magnetic force increases, and eventually becomes large enough to bend the trajectory back until the velocity of the electron again becomes zero and has the same potential in the electric field as that at which it started.

Finally, an electron traveling from right to left (mathematically, a negative velocity) is at first pushed upward by the magnetic and electric fields as you can see in curve *d*. Its initial radius of curvature is small. Soon it goes straight upward gaining velocity rapidly. As the magnetic force increases steadily, the electron is pushed to the right and finally downward. The path is bent back and the electron finally returns to its initial condition.

The time for each cycle and the horizontal distance covered by each electron in one cycle is the same for each of the cases given. All the paths above are cycloids. Electron *a* corresponds to the center of a wheel rolling on a smooth plane. Electron *b* corresponds to a point on the radius, electron *c* to a point on the circumference, and electron *d* to a point outside the wheel.



Cycloids Generated by Points on Radius of a Wheel



Paths of Four Electrons Moving in Mutually Perpendicular Electric and Magnetic Fields

Source of Energy in a Magnetron Oscillator

In an electric field alone, a free electron which always tends to move parallel to the field, gains a large amount of kinetic energy. In combined electric and magnetic fields, the electron, which on the average moves at right angles to the field, does not change its average kinetic energy. The magnetic field always exerts a force which tends to change the direction of motion of the electron.

In a magnetron oscillator there are two components of the electric field: a steady component furnished by the DC source and an alternating component furnished by the oscillations of the tank circuit. In a properly designed magnetron, the electrons absorb energy from the DC component of the field and deliver it to the alternating component. The function of the magnetic field in many respects is the same as that of positive feedback to the grid of the triode oscillator.

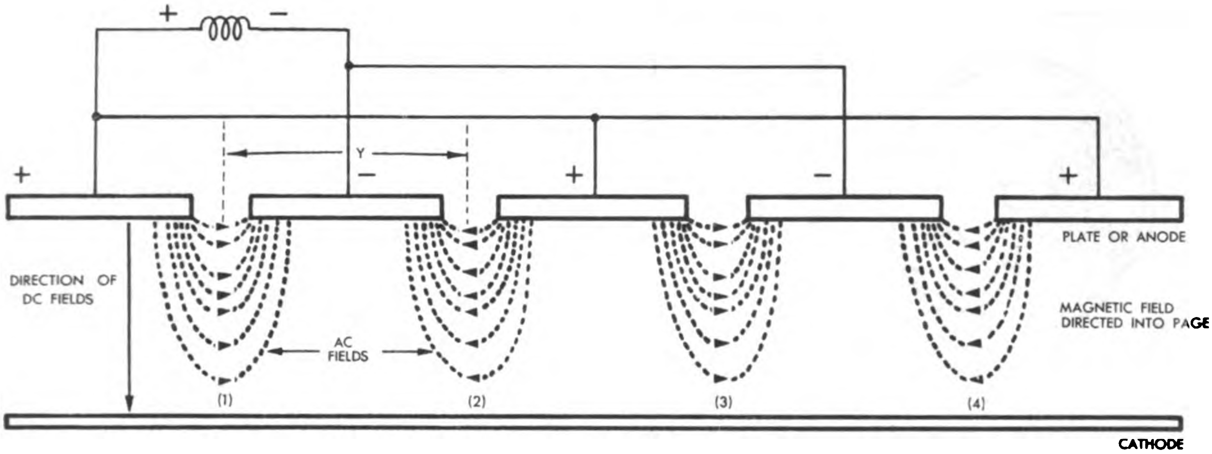
The Multi-Anode Transit-Time Magnetron

Since the theory of the cylindrical multi-anode magnetron is more difficult to understand than that of the plane form, the plane form is discussed first. The ideas presented for the plane form may be carried over in the discussion of the more practical cylindrical multi-anode magnetron.

PLANE FORM. To understand the operation of the plane-form magnetron, consider a magnetron oscillator which is composed of a continuous plane cathode and a segmented anode, the alternate sections of which are connected to the opposite side of its tank circuit. In the illustration at the top of the next page, note that the alternating electric field is sketched in the space between the cathode and anode at the instant when the alternate anode segments are at their maximum positive and negative values.

Let us see which electrons tend to sustain oscillations and which ones absorb energy, and what paths they take. Consider an electron in a uniform electric field directed as shown in the region (1) in the preceding magnetron illustration. The direction of the resultant field is obtained by adding the AC and DC fields at point (1) vectorially. (The DC field, which is not shown, is directed downward.) The approximate direction of the resultant electric field E_1 and the path of the electron are shown in diagram A of the lower illustration on page 345. The electron travels in cycloidal paths. Since the general progression of these cycloids is always at right angles to the resultant electric field, the electrons progress upward and to the right, with the magnetic field directed into the page. On the other hand, an electron in a uniform electric field in region (2) of the magnetron shown will progress downward and to the right as shown in diagram B. The direction of the resultant electric field E_2 is also indicated.

Since the electron of diagram A above moves, on the average, from cathode to anode, it

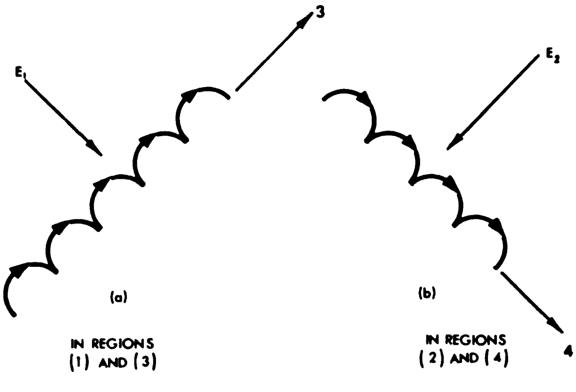


AC Fields in Plane Magnetron

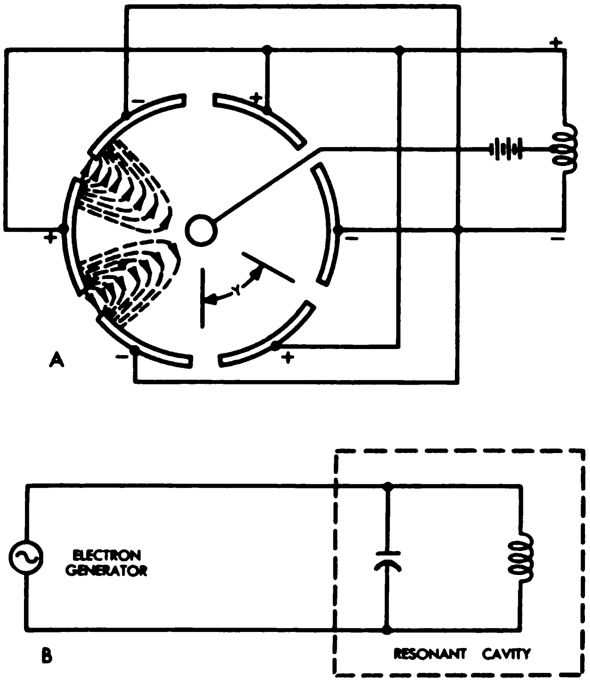
absorbs energy from the steady DC field. This electron, however, tends to deliver energy to the alternating field since it moves to the right in a direction which is opposed by the field of the tank circuit. Because the average velocity V of the electron does not change as it progresses toward the anode, this electron converts energy from the DC source into energy which produces oscillations in the tank circuit. By the same reasoning, the electron in diagram B, which moves downward and to the right, tends to absorb energy from the alternating field, and to transfer it to the steady DC field.

To sustain oscillations, the electrons which follow the path shown in diagram A must be made to continue their travel while the other electrons — the ones which absorb energy — must be removed from the field quickly. An electron of the type in diagram B, starting from the cathode with a small or zero velocity, will strike the cathode before completing its first loop. If the fields shown in regions (1) and (2) are constant, the electron of diagram A will tend to turn downward as it approaches region (2). Since the fields are alternating, the electron of diagram A will encounter in region (2) exactly the same kind of field that it encounters in region (1) if the left-to-right velocity of the electron is such that it arrives in region (2) in exactly one half-cycle. If this electron continues to progress a distance, Y, from left to right during each half-cycle, it will go all the way to the anode, following a cycloidal path similar to that of the electron in diagram A. This condition is necessary for oscillation, since the electron of diagram A contributes energy to the tank circuit for a much longer time than the electron of diagram B takes energy away from the tank circuit. The plane magnetron is a transit-time oscillator, for an electron is allowed a certain definite time to travel from cathode to anode.

CYLINDRICAL FORM. You may consider the cylindrical form magnetron similar to the plane magnetron rolled up into a cylinder. The number of segments (anodes) which may be used in it varies widely. As many as 32 have been tried on certain occasions. Here, however, for simplicity, the discussion considers a magnetron



Possible Electron Paths in a Plane Magnetron



Six-anode Magnetron and Equivalent Circuits

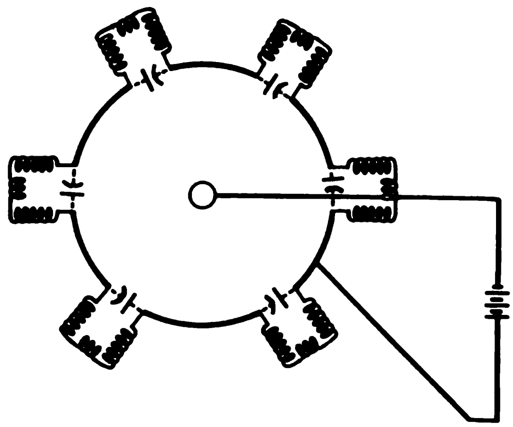
with only six segments. Diagram A above shows the circuit of a six-anode magnetron and the approximate AC field between the segments. The equivalent AC circuit shown in diagram B represents the electrons between anode segments as an electron generator.

As mentioned previously, the exact theory involved in the cylindrical magnetron is more complicated than that used for explaining the plane magnetron. For that reason the discussion does not go into this theory. However, you should qualitatively understand the performance of the cylindrical magnetron. For this purpose, assume that the DC field is uniform and that its magnitude is $\frac{E_b}{b-a}$, where E_b is the DC supply voltage, *b* the anode radius, and *a* the cathode radius in centimeters. Since the AC field is much narrower near the cathode than near the anode as shown at A, the linear distance (the distance which the electron must progress around the cathode per anode segment) is less near the cathode than near the anode. To apply the theory of the plane magnetron to the cylindrical magnetron, you must choose an effective or average value of Y. This value is approximately one-sixth of the circumference of the circle halfway between the cathode and anode.

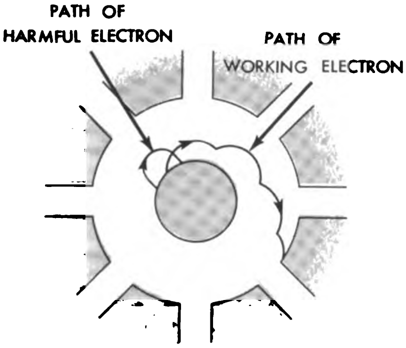
The first illustration on page 347 shows the paths of electrons in a cavity magnetron in cylindrical form. The working electron which is emitted at the proper time and in the proper region to take energy from the DC field and transfer this energy to the AC field makes a number of cycloids as it travels from the cathode to an anode segment. The harmful electron is emitted at a time when it takes energy from the AC field; however, it strikes the cathode and is quickly eliminated. If the magnetron field, the DC potential, and the resonant cavities are properly designed, a cavity magnetron may convert from 40 to 60 per cent of the applied DC power to AC power.

MODES OF OSCILLATION OF THE MAGNETRON. The simplified circuit of the cylindrical magnetron to the left may be split up to form six resonant circuits with magnetic coupling between them as shown below.

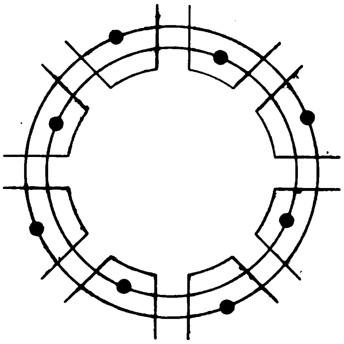
While there is usually considerable coupling between inductances that are close together, in actual magnetrons which employ cavities as resonant circuits there is no coupling between adjacent cavities — that is, coupling in a manner characteristic of lumped circuit constants. On the other hand, there can be coupling resulting from radiation of energy from one cavity to another. As you recall from the theory of coupled circuits, when two circuits tuned to the same frequency are closely coupled, they will resonate not at that frequency but at two different frequencies. These six identical resonant circuits, all closely coupled, resonate, not at two, but at three different frequencies which



Equivalent Circuit of Cylindrical Magnetron



Paths of Electrons in Cylindrical Cavity Magnetron



Strapping in a Magnetron to Obtain a π Mode

differ very little in frequency from each other. For each frequency there is one configuration or mode of oscillation. For example, for three frequencies, there are three configurations as shown below.

The radar magnetron should oscillate only in one mode under operating conditions since the frequencies of these three modes are generally different, and since the radar receiver can be tuned to only one frequency at a time. If the modes are widely separated in frequency, it will be easier to excite only the mode desired. Since mode three requires the least voltage and is the most stable, it is generally used. This mode, in which a phase difference of 180° exists, is sometimes called the π mode. Mode frequency separation may be accomplished by coupling the various cavities more closely by means of conductive connections between them. These connections are known as straps, and the method is known as strapping. A typical example of strapping for the π mode is shown schematically above. The dots indicate connections of the straps to the alternate anode segments.

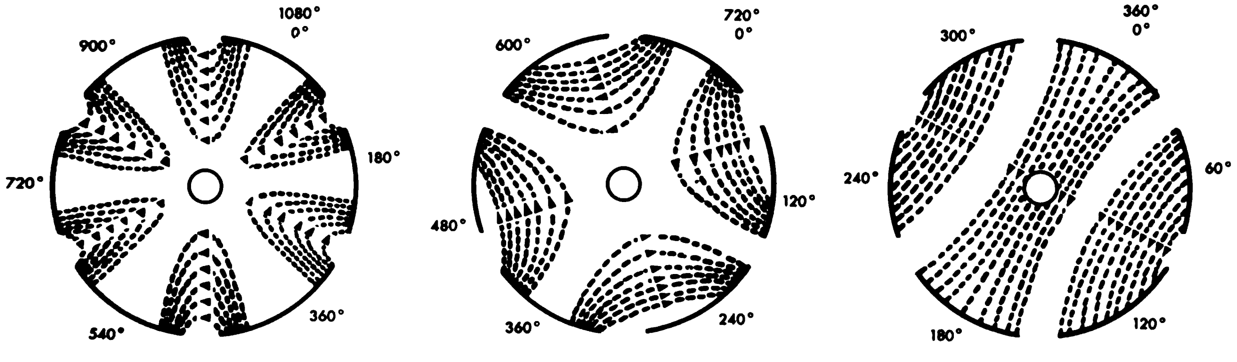
A Review of Important Magnetron Facts

The following summarizes the discussion of magnetrons:

1. By using a segmented anode, you can make the electrons work against an alternating electric field which is crosswise to the steady field. In this way it is possible to get the electrons to go all the way to the anode without much increase in kinetic energy. This results in high power and high efficiency.

2. Oscillations are obtained in the π mode when the average crosswise velocity is such that the electron passes two anode segments per cycle.

3. Magnetrons in use at present have a wavelength range from about 50 cm down to less than one cm.



MODE 3 PHASE DIFFERENCE 180°

MODE 2 PHASE DIFFERENCE 120°

MODE 1 PHASE DIFFERENCE 60°

Three Modes of Operation of a Magnetron

4. Efficiencies of magnetrons at moderate power in service at present have values from 30 to 40 per cent, and from 50 to 60 per cent at peak powers greater than 200 kw.

5. Magnetrons are readily adaptable to pulsed modulation.

REFLEX KLYSTRON

In radar, most receivers use 30 or 60 mc intermediate frequencies. In superheterodyne receiver operation, the stability of the frequency of the local oscillator is a highly important factor. For example, in a radar receiver which receives a frequency of 3000 mc, a frequency shift in the local oscillator as much as 0.1 per cent is a 3-mc frequency shift. This is equal to the bandwidth of most receivers and would cause a considerable signal loss in the receiver.

Many radar receivers use crystal mixers. In this type of receiver the power required of the local oscillator is small, being only 20 to 50 mw in the 3000 mc region. Because of the very loose coupling, only about one milliwatt of the power of the local oscillator actually reaches the crystal.

A third requirement of a local oscillator used in radar is that it must be easily tunable over a range of several megacycles to compensate for changes in the transmitted frequencies and in its own frequency. In cases where the RF units are located in a remote position, the converter section of the receiver may be mounted near the transmitter. Thus the IF rather than the transmitted frequency is cabled back to the receiver. It is therefore often necessary that the local oscillator be tuned by varying the voltage applied to it from a position some distance away.

Because the reflex (velocity-modulated) klystron meets these three requirements better than any other type tube now in production, it is used almost exclusively for local oscillators in microwave radar receivers. The following discussion deals with its operation. In this discussion the resonant cavity is treated as a parallel resonant circuit. You will see the reason for this when you study waveguides and resonant cavities in chapter 36.

A reflex klystron is a tube which uses a resonant cavity as a tank circuit and generates microwave frequencies. Because the various electrons emitted are accelerated or decelerated at different rates (thus varying their velocities), a reflex klystron is known as a velocity-modulated tube.

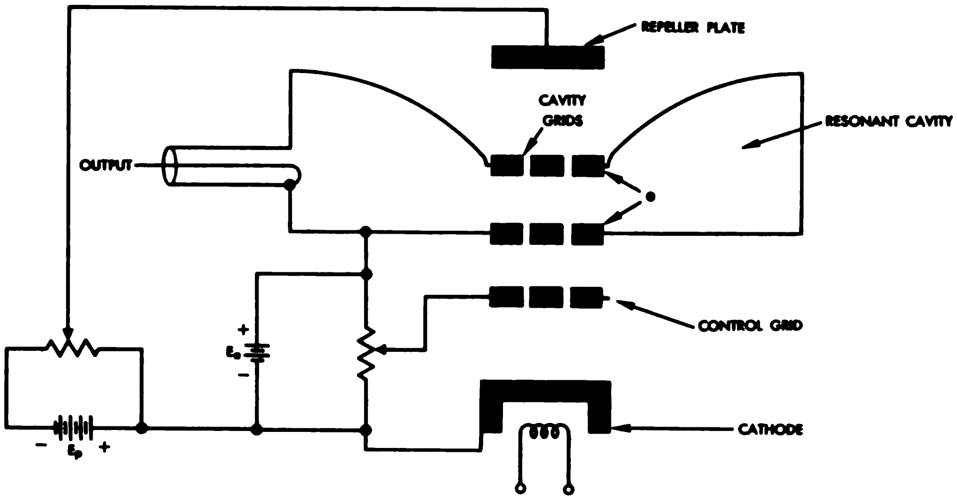
The Path of Electrons in a Klystron

In the circuit of the reflex klystron, note the arrangement of electrodes and the voltages involved for operation. Electrons are emitted by an indirectly heated cathode. These electrons are attracted by the cavity grids which have a positive potential E_a of approximately three hundred volts with respect to the cathode. The control grid which is located between the cathode and the cavity grids is also positive with respect to the cathode. It controls the flow of electrons from the cathode and focuses the electrons into a beam. The electrons emitted from the cathode therefore travel toward the cavity grids at a high velocity. Most of the electrons pass through the control grid, the cavity grids, and continue on toward the repeller plate. After passing the cavity grids, they come to a region where the electrical field opposes their motion since the repeller plate is about 130 to 190 volts negative with respect to the cathode. This negative voltage slows down the electrons causing them to come to a stop, reverse direction, and return through the cavity grids. They are collected either by the control grid, the shell, or the cathode of the tube. The distinguishing feature of this tube is the repeller plate, or reflector, which gives the *reflex* klystron its name.

Bunching Electrons

An important consideration is the cause of oscillations in resonant cavities. In most oscillators, oscillations start from some irregularity in current flow, such as a transient that results from voltage being suddenly applied to the tube. With this in mind, assume that oscillations in the resonant cavity are already taking place. From this assumption, examine the source of the energy needed to sustain these oscillations. With the circuit oscillating, a high frequency voltage, e, appears between the two cavity grids. This makes the electric field between these grids reverse twice each complete cycle of operation.

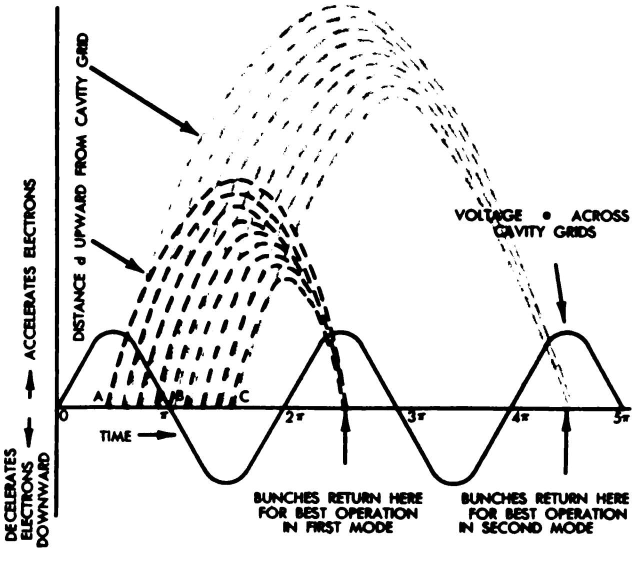
As the electrons approach the cavity grids, the electron stream is uniform. The time that is



Reflex Klystron Circuit

required for the electrons to pass through the small distance between the grids is small compared to the period of oscillations. Electrons which enter the space between the grids when e is zero encounter no AC electrical field and they pass on through at the same velocity. The electrons which enter the space when e makes the top grid positive with respect to the bottom will encounter a field which tends to accelerate them. The amount which they are accelerated depends upon the amplitude of e. Electrons entering the space when e is reversed in polarity are decelerated. The change in velocity due to acceleration and deceleration is small compared to the original velocity. Those electrons which are accelerated most will travel farther toward the repeller plate before being turned back, while those that are decelerated most will be turned back before approaching very close to the repeller. Therefore, electrons returning to the cavity grids arrive in bunches.

The illustration to the right shows the position of electrons in the tube at various times during their transit. The zero distance position is midway between the cavity grids. Electron A, which arrives when e is positive, is accelerated and travels farther toward the repeller plate before being turned back. Electron B is unaffected, and electron C is decelerated and turned back after traveling a shorter distance. In the diagram these electrons and the ones passing through at intermediate times are shown arriving back at the grids at the same instant. This is the ideal situation. Since electrons are emitted in a constant stream, it is not difficult to see that electrons will return to these grids in a stream which varies in intensity at the frequency of the oscillations. It is due to this fact that this tube is called a velocity-modulated tube. It is called a reflex velocity-modulated tube because the electrons reverse direction and travel through the interelectrode space twice.


DISTANCE d UPWARD FROM CAVITY GRID
ACCELERATES ELECTRONS
DECELERATES ELECTRONS DOWNWARD
VOLTAGE e ACROSS CAVITY GRIDS
A
B
C
0
τ
2τ
3τ
4τ
5τ
TIME
BUNCHES RETURN HERE FOR BEST OPERATION IN FIRST MODE
BUNCHES RETURN HERE FOR BEST OPERATION IN SECOND MODE

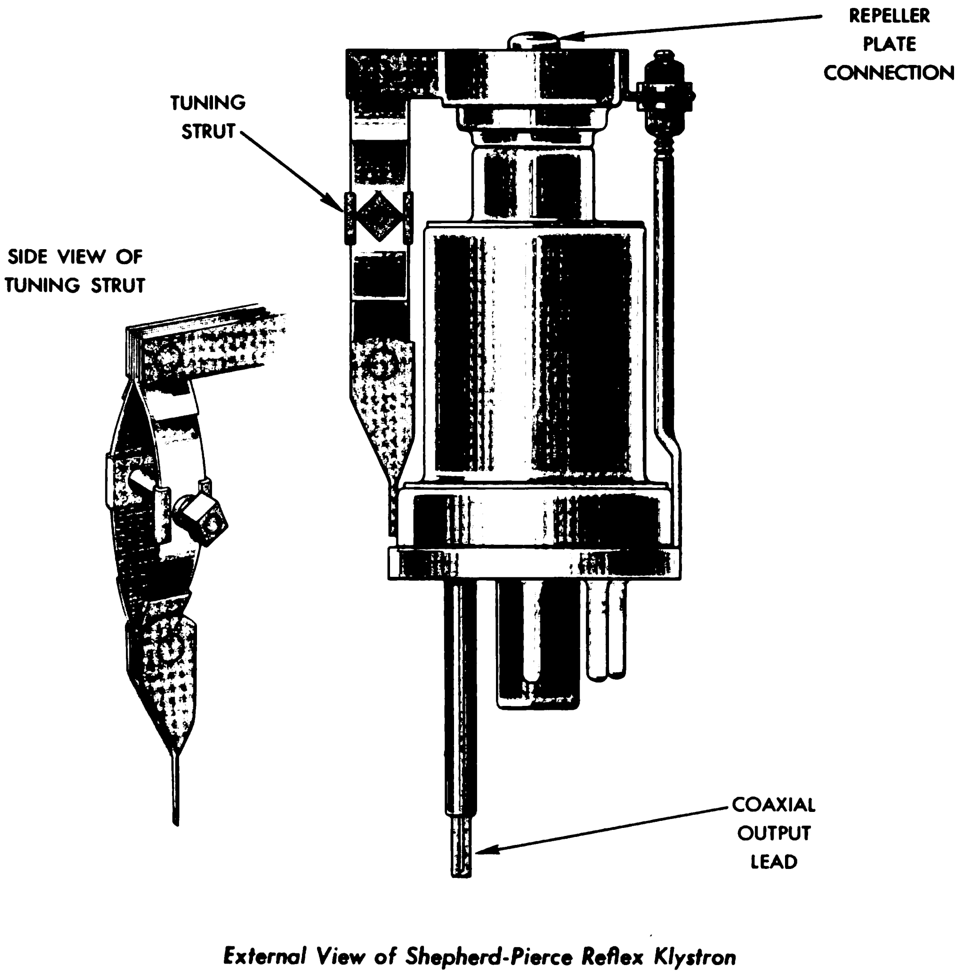
Bunching Action of Reflex Klystron

Sustaining Oscillations

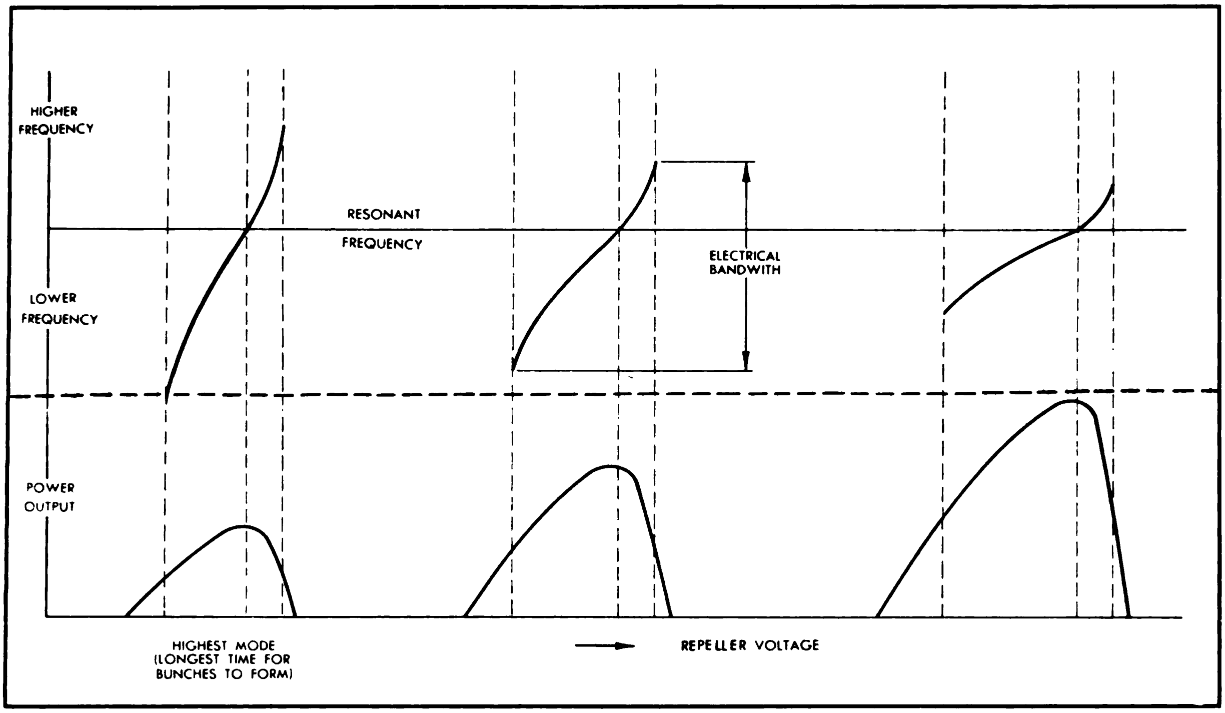
On the return trip the electric fields set up by the voltage e again act upon the electrons. Since they are now traveling in the opposite direction, they will be decelerated if they return when e is positive, and accelerated if they arrive when e is negative. You are already familiar with the fact that an electron which is accelerated by an electric field has its kinetic energy increased and that this additional energy is taken from the electric field. On the other hand, an electron which is decelerated gives up energy to the electric field. Now, if the bunches of electrons can be made to arrive back at the grid when e is positive, they will give up energy to the alternating field. For maximum transfer of energy, the bunches must arrive when e is maximum positive. The question arises as to where did this energy originate. You have seen that if the electron stream from the cathode is uniform, some electrons are accelerated and some are decelerated, on the outbound trip, by the electric field of e. On the average, as many electrons absorb energy from the field as give up energy to it. Hence, very little net energy is taken from the oscillating circuit during the bunching process. The average kinetic energy of the electron is that imparted to it by the DC voltage. Thus, energy is taken from the DC electric field and transferred to the AC field, and oscillations are sustained.

Modes of Operation

It is not necessary that the bunches of electrons return to the grids on the first positive swing of e after they leave them. The illustration of bunching on page 349 indicates that if the bunches arrive on the second positive swing, the net result is still the same. You can see



External View of Shepherd-Pierce Reflex Klystron



Power Output and Frequency versus Repeller Voltage

that the time in transit for the average electron B may be ¾ cycle, 1¾ cycles, 2¾ cycles, etc. In actual practice there are three of four *modes* in which it is possible for the reflex-velocity — modulated tube to oscillate. The black lines in the illustration indicate paths of electrons operating in the first mode while the gray lines indicate paths to electrons operating in the second mode. A third mode is possible when the average transit time is 2¾ cycles, etc.

The transit time may be controlled to produce oscillations in the different modes. Since the original velocity of an electron depends on the DC voltage E_a, and since the distance that the electron travels before turning back and the speed with which it returns depend upon the difference between E_a and E_r, it is possible to adjust the two voltages E_a and E_r to operate in any of the modes. The voltage E_a is usually fixed, since varying it produces a greater initial velocity. However, E_r is usually variable. For operation in the first mode, the round trip must be completed in the shortest time. This is accomplished by making the repeller plate most negative. For operating in other modes the repeller is made less negative.

In the above diagram showing power output and frequency of oscillations as functions of the repeller voltage for three modes of operation, notice that the frequency at the point of maximum output is the same for all three modes and is the resonant frequency of the cavity. In addition, note that the power output for the various modes at the resonant frequency is not the same and the output is least in the highest mode.

The band of frequencies which can be obtained by varying the repeller voltage lies between the half power points shown in the diagram above. This range of frequencies is known as the *electrical bandwidth.*

The choice of the point and mode of operation is a compromise among several factors. To begin with, there are three or four modes which have the necessary power output. On the whole, it would seem then that the correct choice would be the highest mode, for it gives the largest tuning range. The highest mode, however, is too sensitive to a change in voltage to be very well regulated. A change of one volt may cause a change of 0.5 mc in the 3000 mc oscillator. Since the humps are unsymmetrical, the point of operation is usually chosen a little below the point of maximum output. This makes possible the tuning above the operating frequency by a greater amount than if the maximum point were used.

In practice, the reflex velocity-modulated tube is usually used with an automatic frequency control circuit which controls the repeller voltage in such a way as to maintain the correct intermediate frequency. Keep in mind that the frequency of oscillations is primarily determined by the dimensions of the cavity and that the repeller voltage is effective only in making small changes in the frequency. In most reflex velocity-modulated tubes there is a coarse frequency adjustment which varies the cavity size in some way, and the repeller voltage is the fine frequency adjustment. In the illustration on page 350 is an external view of a Shepherd-Pierce reflex klystron showing its tuning strut. This tuning strut is a coarse frequency control which effectively varies the distance between the cavity grids and tunes the cavity by changing the capacitance of the resonant cavity.

Summary of Characteristics of a Reflex Klystron

The more important characteristics of a reflex klystron may be summarized as follows:

1. A reflex klystron has several modes of oscillation, each of which may be obtained by varying the repeller plate voltage sufficiently to allow the oscillator to go in and out of oscillation.

2. As the repeller plate voltage is increased within a mode, (a) the power output increases from zero to a maximum and then decreases to zero again; (b) the frequency of operation varies from below resonance, to resonance, to above resonance; and (c) the maximum power output and the resonant frequency occur for the same value of repeller plate voltage.

3. Of the modes obtainable for a given reflex klystron, the highest mode has the least power output and the greatest electrical bandwidth.

4. The frequency of operation of a reflex klystron is primarily determined by the resonant frequency of the cavity resonator but can be controlled over a range of approximately one per cent as the repeller plate voltage is varied. The resonant frequency of the cavity may be varied mechanically.

5. The output of a reflex klystron is taken from the cavity by means of a loop or a probe through a coaxial lead. This coaxial lead may then be coupled to a coaxial line or to a waveguide, depending upon which type of transmission line is being used.

6. A reflex klystron is used as a local oscillator in most radar receivers operating at frequencies above 2500 megacycles, owing to the fact that the reflex klystron has the following desirable characteristics of a local oscillator: (a) at least 20 milliwatts of useful power is available. (b) it has good frequency stability, (c) it is possible to control the frequency of the oscillator by varying the repeller plate voltage by means of a tuning control near the operator. This may also be accomplished automatically through the use of an automatic-frequency-control circuit.

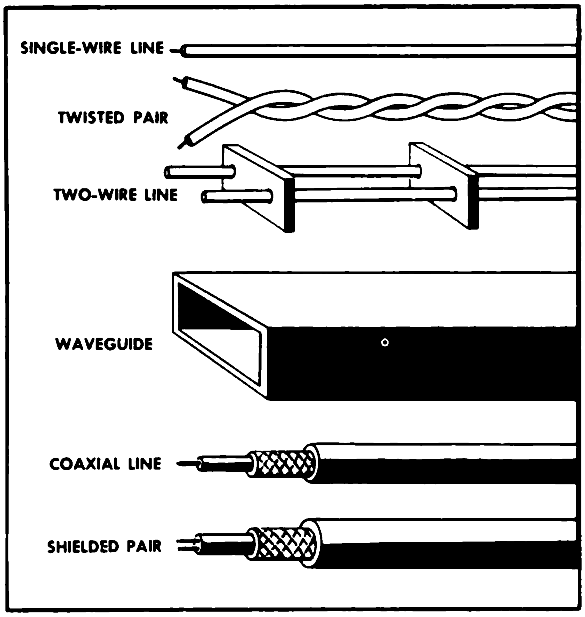
Chapter 35

Transmission Lines

You make use of a transmission device each time you drive your car. The transmission transfers the power of the engine to the rear wheels; it is the connecting link. In radio or radar we must have a connecting link between the transmitter and antenna, or between the receiver and antenna. This link is called a transmission line. The transmission line transfers energy from the transmitter to the antenna or energy from the antenna to the receiver. If the antenna could be built directly on the transmitter and the receiver, then there would be no need for the transmission line. However, even in a compact installation, such as an aircraft, it is not usually convenient to locate the transmitter or receiver near the antenna. Perhaps you have gained the impression that transmission lines are used only for conducting radio frequency energy from one point to another. This is not true. The lines that supply electrical energy to your house are transmission lines just as much as are the wires connecting one circuit to another.

A transmission line can therefore be defined as a device for transmitting or guiding electrical energy from one point to another. A poor transmission line decreases the amount of energy available at the transmitting antenna for radiation or at the receiver for reception. Conversely, a perfect transmission line delivers the entire transmitter output to the antenna or the entire received signal to the receiver. In such an ideal system, the antenna alone radiates or receives energy.

RF transmission lines behave differently in different frequency ranges. An RF transmission line designed to operate at the relatively low frequencies used in the standard broadcast range is not likely to be usable in point-to-point aircraft communication systems. An RF transmission line can be expected to transfer energy efficiently from one point to another only at the frequency for which it is designed. When this line is used at higher frequencies, it will have greater losses. In practice this means that



Types of Transmission Lines

different kinds of transmission lines must be available for use in different frequency ranges.

In this chapter, the basic types of RF transmission lines are described. The characteristics of each type are discussed in order to determine how the line operates and how energy is lost. Methods of obtaining maximum efficiency of power transfer are also described.

BASIC TRANSMISSION LINES

The job of the transmission line is to get the energy from the place where it is produced to the place where it is to be used and to accomplish this with a minimum amount of loss. In a transmitter the transmission line usually runs between the final tank circuit of the last power amplifier stage to the transmitting antenna. In a receiver, the transmission runs from the receiving antenna to the first tuned circuit of the receiver.

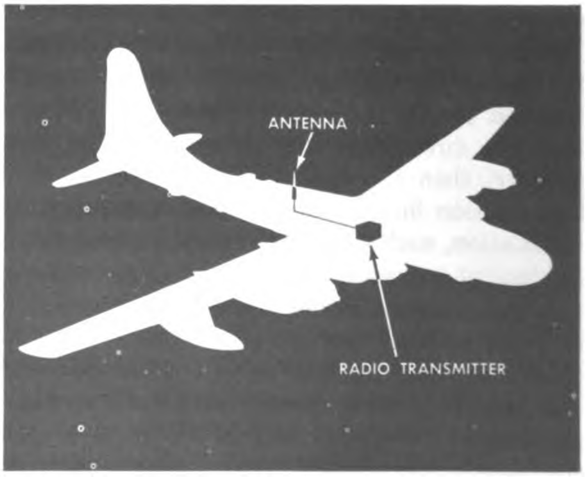
In some applications, the transmission line is only a few feet long. In certain aircraft radar equipments, for example, the transmitter is located quite close to the radiating element, and only a very short run of transmission line is needed. On some fixed land installations, the transmission line may be several hundred feet or even longer.

There are many types of transmission lines commonly used today. Some of these are shown in the illustration to the left. Each type of transmission line has certain advantages and disadvantages compared to the other types. Certain types are more suitable at UHF or SHF than others. Waveguides, for example, are commonly used for these higher frequencies. Each type of transmission line is discussed more fully later in this chapter.

Terminology

All transmission lines have two ends. The end of the line where the energy is injected is known as the *input end.* This end is also referred to as the transmitter end, the sending end, the generator end, or the source. The end from which the energy is taken and used is termed the *output end.* This end is also known as the receiving end, or the sink. The input and output ends are shown on the simple transmission line in the illustration below. In this case, the line is being used to couple the output of the transmitter to an antenna mounted on the fuselage of an aircraft.

You can describe a transmission line in terms of impedance. For example, the ratio of voltage to current at the input end is termed the *input impedance.* This is the impedance presented to the source by the transmission line and its load (the antenna). Similarly, the ratio of voltage to current at the output end is called the *output impedance.* This is the impedance presented to the load by the line and its source. In the case shown, the source is the radio transmitter. In other words, the input impedance is that impe-



Simple Transmission Line in Aircraft

dance seen at the input end looking down the line toward the load. In a similar manner, the output impedance is that impedance seen at the output end looking down the line toward the source. If an infinitely long transmission line (one that has no end) could be used, then the ratio of voltage to current at any point on that line would be some definite value of impedance. This impedance is known as the *characteristic* or *surge* impedance of the line. The characteristic impedance is discussed fully later in the chapter.

Wavelength on a Transmission Line

You have been told previously that radio-frequency energy travels with the speed of light, approximately 186,000 miles per second. However, this is true only in free space. In a transmission line, factors in the construction of the line modify this velocity of propagation. The ratio of the velocity of RF energy in a given line to the velocity of RF energy in free space is known as the *velocity factor*. The symbol for the velocity factor is K. This factor is dependent on the characteristics of the dielectric material used for the spacing of the line and/or supporting insulators used in the construction of the line.

If an electrical wavelength is measured on a transmission line, it usually will be different from a wavelength at the same frequency measured in free space. This is so because the wavelength is equal to the velocity of propagation divided by the frequency. Since the frequency of the RF energy remains constant whether the wave is traveling in free space or on a conductor, the wavelength must change when the velocity of propagation varies. The velocity factor of the specific transmission line shows the amount by which a wavelength measured on the line will be different from a wavelength at the same frequency measured in free space. Velocity factors for various types of basic transmission lines are shown in the table.

VELOCITY FACTORS

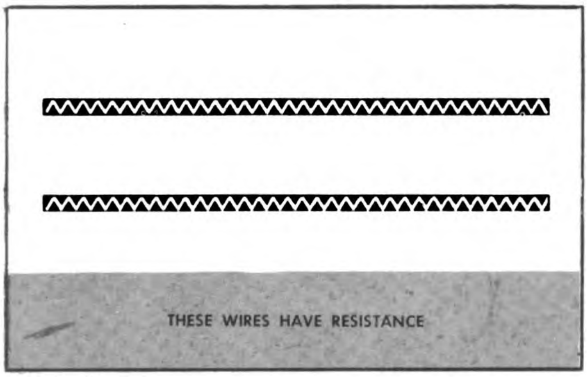
Type of Line	K
Two-wire open line (wire with air dielectric)	0.975
Parallel tubing (air dielectric)	0.95
Coaxial line (air dielectric)	0.85
Coaxial line (solid plastic dielectric)	0.66
Two-wire line (wire with plastic dielectric)	0.68-0.82
Twisted-pair line (rubber dielectric)	0.56-0.65

Properties of Transmission Lines

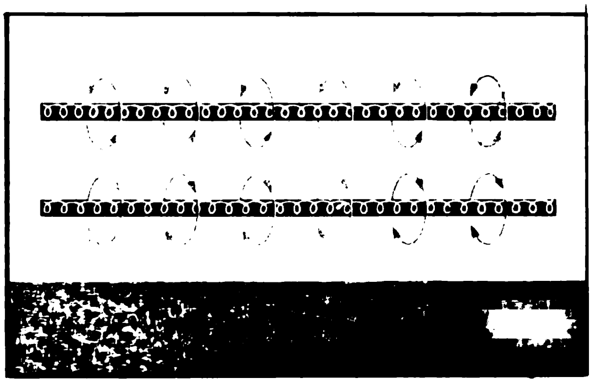
A transmission line is considered to be short when its physical length is small compared to a quarter wavelength of the energy which it is to carry. A line is long when its length is large compared to a quarter wavelength. It is important to understand that the term *short* and *long* are relative ones. For example, a line whose physical length is 10 feet would be considered quite short if it were to carry radio frequencies in the standard broadcast band (from 550 to 1,600 kc). On the other hand, if this same length of transmission line were to be used to carry ultrahigh frequencies, the line would be very long indeed.

When power is applied to a very short transmission line, practically all of that power gets to the load at the output end of the line. This very short transmission line is usually considered to have practically no electrical properties of its own except perhaps for a small amount of resistance. However, the picture changes considerably when a long transmission line is used. As most transmission lines to be considered in this manual are long, it is of interest to consider the properties of such a line. There is a reason for doing this. It frequently happens that the voltage necessary to drive a current through such a long line is sometimes considerably greater than can be accounted for by the impedance of the load in series with the resistance of the line.

DISTRIBUTED CONSTANTS. A transmission line has all the properties of more conventional circuits, such as inductance, capacitance, and resistance. In most cases, however, the constants in conventional circuits are lumped into a single device or component. For example, a coil of wire has the property of inductance, so that when



Distributed Resistance



Distributed Inductance

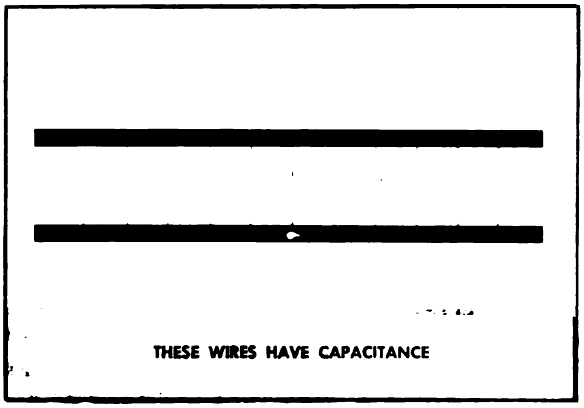
a certain amount of inductance is needed in a circuit, a coil of certain dimensions can be inserted. The circuit's entire inductance is lumped into the one component. Two metal plates separated by a small space can be used to supply the circuit capacitance. In such a case, the circuit's entire capacitance is lumped into this one component. Similarly, a fixed resistor can be used to supply a certain value of circuit resistance.

A transmission line also has the properties of inductance, capacitance, and resistance. However, these are not lumped. Instead, they are called *distributed* constants; that is, they are spread along the entire length of the transmission line and cannot be distinguished separately. A typical two-wire transmission line, for example, has electrical resistance all along its length. This resistance is usually expressed in ohms per foot. In the illustration, Distributed Resistance, this resistance is shown as existing continuously from one end of the line to the other.

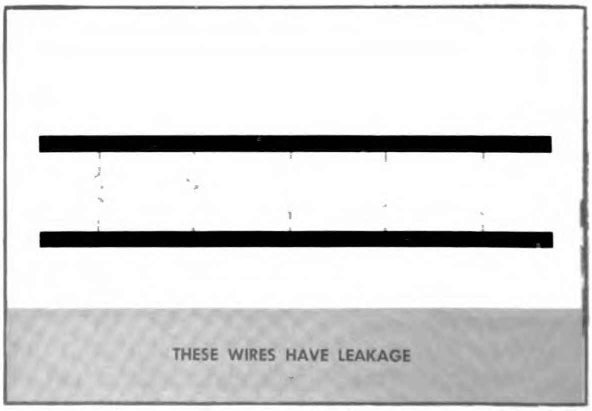
These same two wires also possess inductance. Any current in the wires sets up magnetic lines of force which encircle the wires as shown in the illustration entitled Distributed Inductance. When the current tends to drop, the field collapses and a certain amount of energy is applied to the wires. This energy tends to keep the current flowing a little longer in the same direction. This property represents a certain amount of inductance which is expressed in microhenries per foot.

An electric field similar to the field that exists between the two plates of a capacitor also exists between the two wires. In other words, the two wires of the line act as a capacitor, with air or other insulating material acting as the dielectric between them. This is shown in the illustration, Distributed Capacitance. The capacitance between the wires is usually expressed in micromicrofarads per foot.

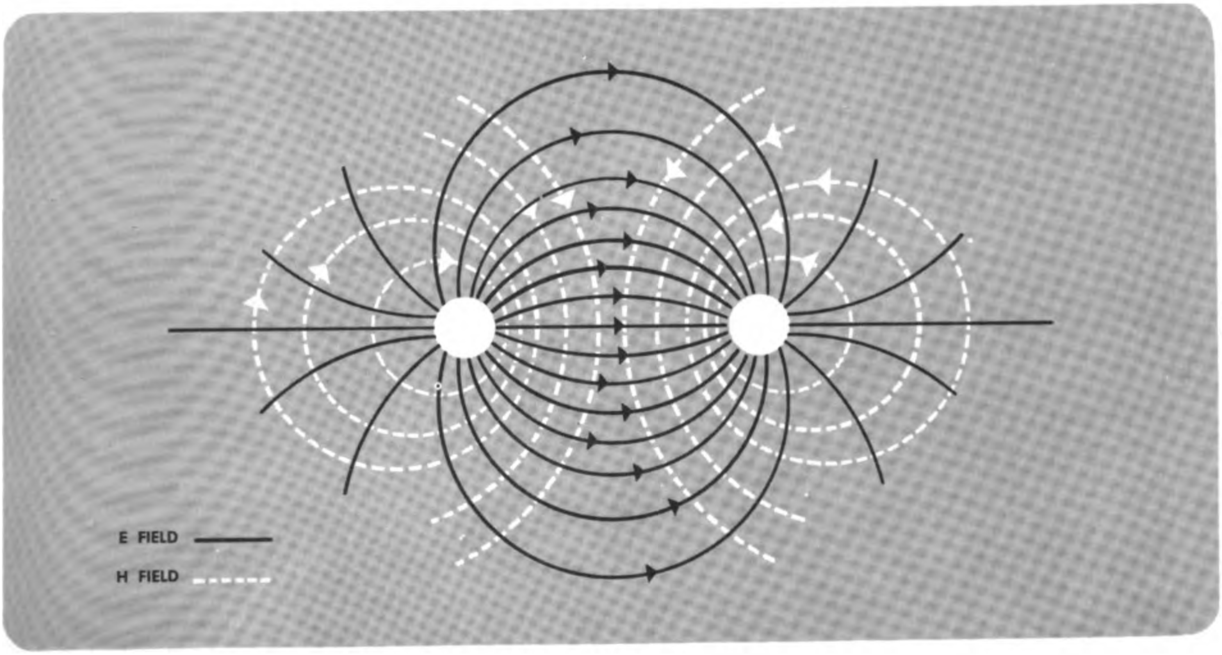
Since any insulator, even air, is not perfect, there is a small current which flows between the



Distributed Capacitance



Leakage in a Transmission Line



Fields between Conductors

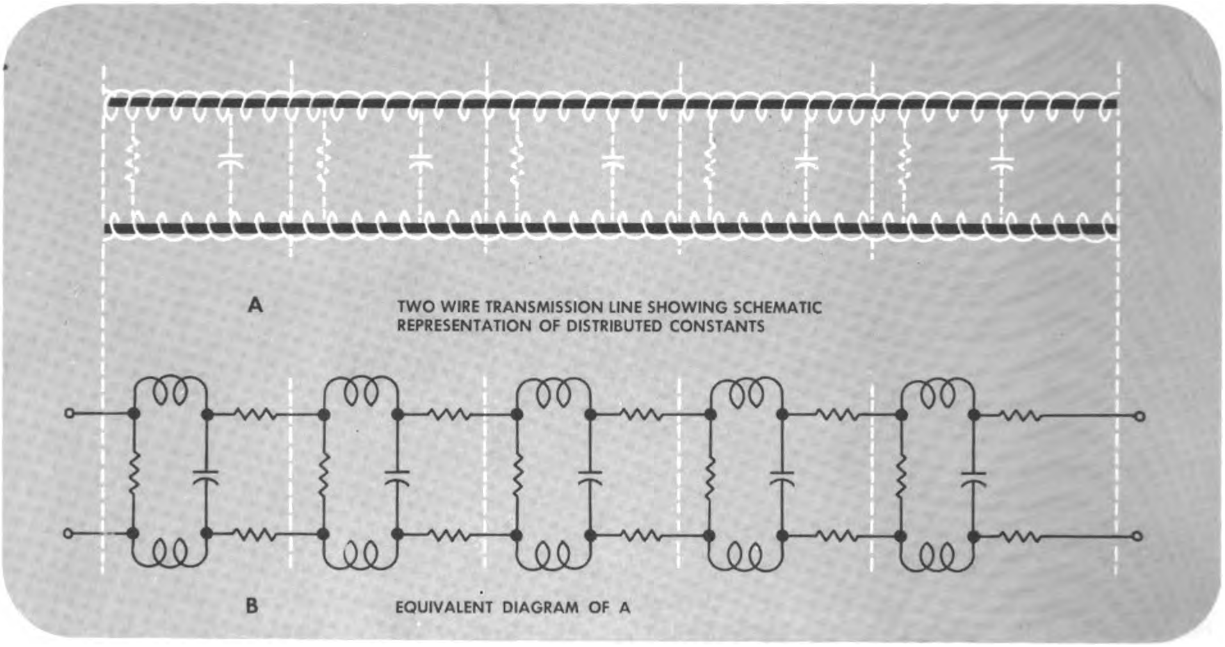
two wires known as leakage current. In effect, the insulator acts as a resistor, permitting current to pass between the wires. This is shown on page 356 in the illustration, Leakage in a Transmission Line. This property is usually expressed as the reciprocal of resistance (conductance) for ease in computation. Conductance (G) in transmission lines is often expressed in micromicromhos per foot.

ELECTROMAGNETIC FIELDS. The distributed constants (resistance, inductance, and capacitance) are basic properties of all transmission lines and exist whether there is any current flowing or not. As soon as current flows and voltage exists in a transmission line, another property becomes evident. This property is the existence of a field, or lines of force, about the conductors which make up the transmission line. While the lines of force themselves are imaginary, the force an electron experiences while in their vicinity is very real indeed.

Actually, there are two kinds of fields, one of which is associated with the voltage and the other with the current. The field associated with the voltage is called the electric or E-field, because it tends to exert a force on any electric charge (such as an electron or an ion) placed in its vicinity. The field associated with the current is called a magnetic or H-field, because it tends to exert a force on any magnetic pole (such as is associated with a magnet or a wire with current flowing in it) placed in its vicinity. The illustration above, Fields between Conductors, shows the way in which the E-field and H-field tend to orient themselves between conductors of a typical two-wire transmission line. The transmission line is shown in cross-section with the E field represented by the solid lines and the H field by the dashed lines. The arrows indicate the direction of the lines of force. Both fields usually exist together and are spoken of collectively as the electromagnetic field.

CHARACTERISTIC IMPEDANCE. Every transmission line, in addition to the properties already mentioned, also possesses a certain characteristic impedance, usually designated as Z_o. The characteristic impedance determines the amount of current that can flow when a given voltage is applied to an infinitely long line. It has much the same effect as resistance has in a DC circuit in determining the amount of current that can flow when a given voltage is applied.

To understand the significance of Z_o, take the equivalent circuit of an infinitely long transmission line as shown on page 358. Assume that all inductors have the same value and that all capacitors have the same value. If an electrical impulse is applied to the input end of this line,


A
TWO WIRE TRANSMISSION LINE SHOWING SCHEMATIC REPRESENTATION OF DISTRIBUTED CONSTANTS
B
EQUIVALENT DIAGRAM OF A

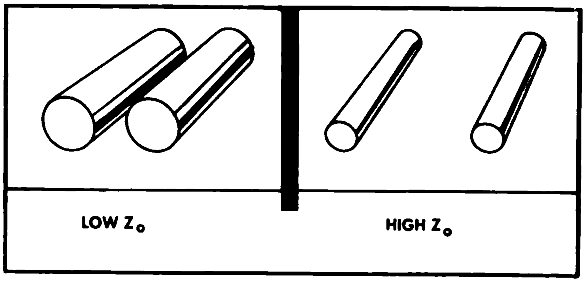
Equivalent Circuit of a Transmission Line

the combination of L and C appears to have an impedance Z_0. This impedance, called the characteristic impedance, is approximately equal to $\sqrt{L/C}$, where L and C are the inductance and capacitance, respectively, per unit length of the line. Z_0 turns out to be purely resistive, and it is a constant value for a given transmission line.

The inductance and capacitance per unit length of line depend upon the size of the conductors and the spacing between them. The closer the two conductors are to each other and the greater their diameter, the higher is the capacitance and the lower is the inductance. The characteristic impedance will therefore be lower. As is shown in the illustration to the right, a line with large conductors closely spaced will have a low Z_0, while one with small conductors widely spaced will have a relatively high Z_0.

When a line is terminated in a resistance equal to its characteristic impedance, the line acts as if it is infinitely long. The characteristic impedance of lines used in actual practice usually lies between 50 and 600 ohms.

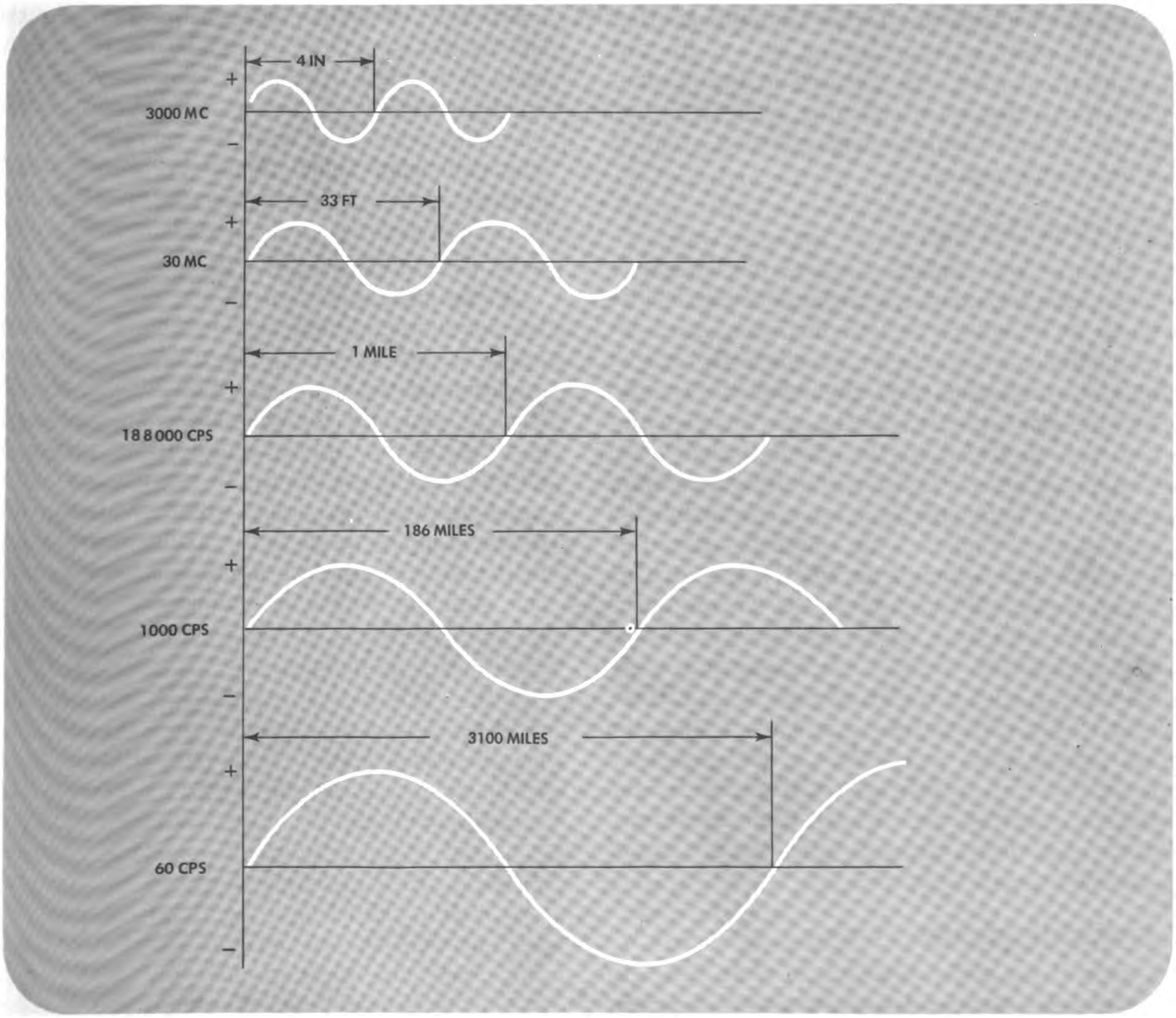
ATTENUATION AND LOSSES. In the introduction to this section on transmission lines, you learned that a perfect transmission line delivers the entire transmitter output to the antenna, or the entire received signal to the receiver. In this ideal situation, no power is lost in the line. In practice, however, there is always some loss in transmission lines. This loss, or *attenuation*, of power may be due to three causes: radiation, heating (I^2R loss), and reflection (in which power is returned to the source). There is no appreciable radiation in coaxial lines, while radiation losses



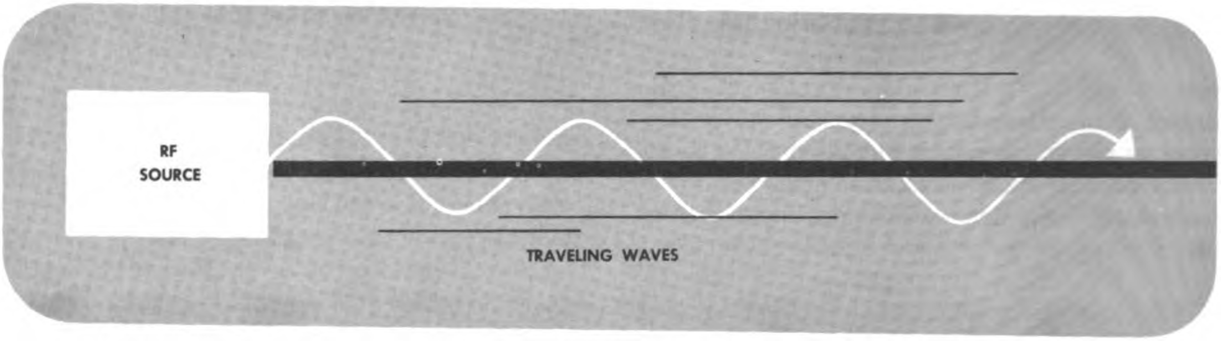
Effect of Conductor Size and Spacing on Characteristic Impedance may be considerable in an open-wire line. Heat losses in both conductors (due to resistance and skin effect) and dielectric losses (resulting from imperfect insulation of the transmission line) increase as the frequency is increased. I^2R losses in conductors are also greater in lines of lower characteristic impedance, because the current which can flow for a given voltage at the input end is greater. An understanding of how reflection losses occur requires an understanding of wave motion on transmission lines. Study the discussion of the motion of waves on transmission lines which follows.

WAVE MOTION ON TRANSMISSION LINES

Waves exist on a transmission line because a certain amount of time is required for electrons to transfer their energy down a wire. At radio frequencies, it is possible that a number of waves will appear on a transmission line of practical length, as shown below in the illustration, Waves on Conductors. For example, at 30 mc the distance between positive voltage (or current) peaks on the line is about 33 feet. If a transmitter producing 3,000-mc RF energy is connected to the



Waves on Conductors



Wave Motion on an Infinite Line

line, the distance between peaks will be 4 inches. In other words, the transmitter puts a positive voltage peak into the line, the peak moves along the line, and four inches behind it another positive peak is formed. The waves are not stationary as they are pictured in the illustration, but actually are moving along the line, much as water ripples move when a stone is dropped into a pool.

Travelling Waves

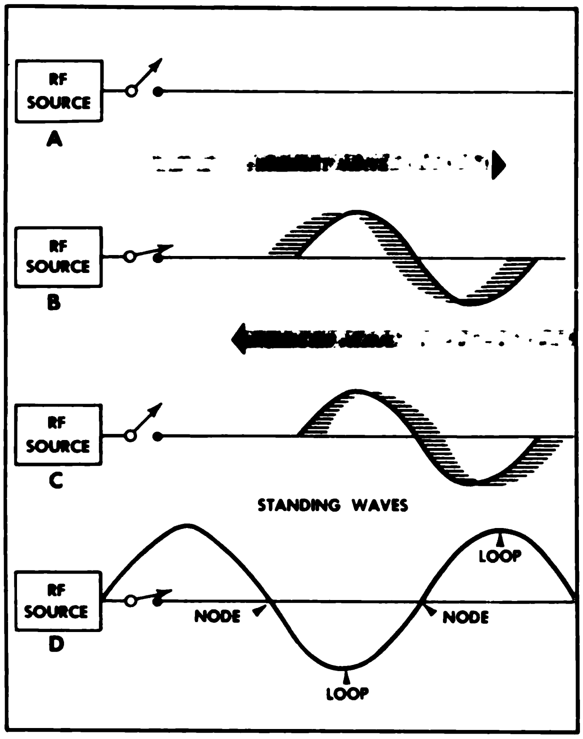
Suppose you connect a source of RF energy to an infinitely long conductor as shown in the illustration above. A current, with its associated field, begins to move down the length of that conductor in the form of waves of energy. As this line never ends, the waves of energy continue to travel in the direction indicated. The energy moving in this manner is appropriately referred to as traveling waves. Actually, the amplitude of the traveling waves is reduced as the wave moves away from the RF source. This occurs because the long conductor has resistance which absorbs some of the energy as it travels. In practice, there is no such thing as an infinitely long line, as such a line would have no end. However, it is possible to present an impedance to the RF source which will make the energy appear to be traveling on such a line. How this is done will be described later. Under these conditions, the energy moves along the line in the form of traveling waves.

Producing Standing Waves

Now assume that a conductor of finite length is connected through a switch to an RF generator as shown in part A of the illustration on page 361. If the switch is closed for a time equal to the period of the wave of energy and then opened, one cycle of energy is sent down the line to the far end (see part B). This wave is called an *incident* wave. When the energy reaches the end of the conductor, it can go no farther as there is no path for current flow beyond this point. As a result, the energy stops abruptly, and the fields associated with the energy collapse. The collapsing fields induce a voltage which causes current to flow back toward the RF source. Thus, a *reflected* wave is caused to move back along the conductor (see part C).

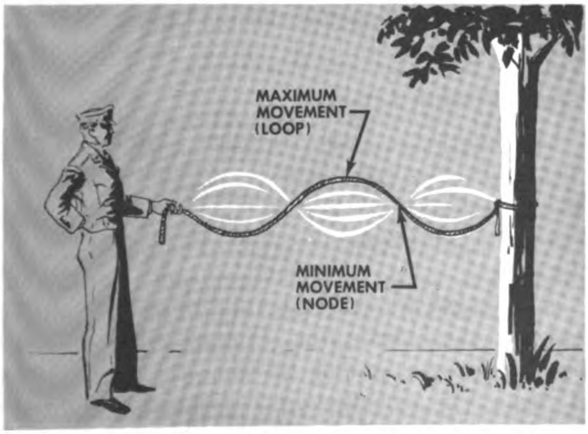
If the switch connecting the RF source to the conductor is kept closed, then a condition exists in which energy is traveling in both directions at the same time. You will see shortly that, if you add together the incident waves and the reflected waves, the resultant waves do not travel. These resultant waves are called *standing waves*. The points where the amplitude of the standing wave is a maximum are called loops; while the points where the amplitude of the standing waves are minimum are called nodes.

Although standing waves on conductors are invisible, it is possible to see standing waves in mechanical systems. If you tie one end of a length of rope to an upright stanchion or post, as shown in the lower illustration on page 361, you can set up standing waves quite easily. Move your end of the rope up and down just once and watch the mechanical wave which you have set up. Watch this wave move along the rope, hit the end tied to the stanchion, reverse its direction, and come back toward you. You will be able to feel the energy come back to you. Now move your end of the rope up and down rapidly and watch what happens. If you shake the end at



Wave Motion on a Finite Line

the right frequency, standing waves are set up. You will note that the rope vibrates in sections — portions of the rope oscillate violently while other portions hardly move at all. Those portions of the rope where maximum movement occurs are the loops of the standing waves, while those portions where little or no movement occurs are the nodes of the standing waves. These waves have been produced by mechanical waves moving in opposite directions through a single conductor.



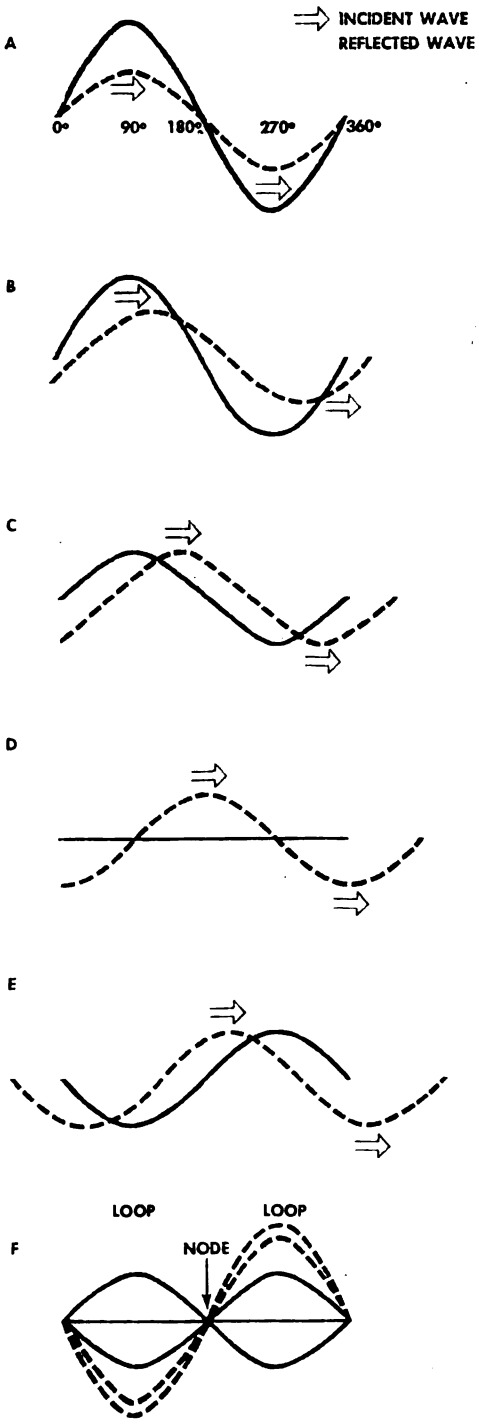
Standing Waves on a Rope

The series of diagrams shown on page 362 in the illustration, Production of Standing Waves, shows how two waves of equal frequency and amplitude moving in opposite directions on the same conductor will combine to form a standing wave. The broken-line waveform is moving steadily from left to right and is considered to be the incident wave. The white waveform is moving steadily from right to left and is considered to be the reflected wave. In part A, the two waves are in phase. The resultant waveform (shown by the solid black line) has a peak amplitude which is equal to the sum of the peak amplitudes of the incident and reflected waves. This resultant is in phase with its two components.

In part B, the incident wave has moved 30° to the right, a distance equal to one-twelfth of a wavelength. At the same time, the reflected wave has moved to the left by the same amount. When these two waves are added together point-by-point and the resultant is obtained, you can see that the amplitude of the resultant is less than that in part A, and the resultant peak occurs midway between the peaks of the component waves. Note also that the resultant wave is in phase with the resultant in part A. The points of minimum and maximum amplitude of both resultants coincide although the amplitudes of both resultant waves are not equal. In part C, both incident and reflected waves have moved 60° or one-sixth of a wavelength, from their original positions. Another reduction in the amplitude of the resultant occurs, but the position of this resultant is exactly the same as in parts A and B.

In part D, the incident and reflected waves have moved 90°, or one-fourth of a wavelength, from their original positions. The two waves are now 180° out of phase with each other. The resultant is thus zero all along the conductor. In part E the two waves have moved a total distance of 120° or one-third of a wavelength, from their original positions. The polarity of the resultant wave has now reversed but note that the points of minimum and maximum amplitude still occur at precisely the same places along the conductor.

Part F shows the five resultant waves obtained in parts A through E. In addition, the two resultant waves, which would be obtained when the incident and reflected waves have moved respectively 150° and 180° from their original



Production of Standing Waves

positions, are shown dashed. Although the amplitudes of these resultants change, these waves do not move along the conductor. These are standing waves. The points of maximum amplitude (loops) coincide and the points of zero amplitude (nodes) coincide.

Impedance of the Line

The impedance of a line at any point is equal to the voltage at that point divided by the current at that point. With DC or at low frequencies, the ratio of voltage to current is the same at all points on the line. The impedance of the line is therefore constant. This is also true of lines carrying RF energy which are terminated in their characteristic impedances (Z_o). However, if there are reflections in the line, the ratio of voltage to current varies along the line. This happens because the voltage and current are not in phase. As you will see, the impedance at one point on the line may be zero, while at another point the impedance may be infinite.

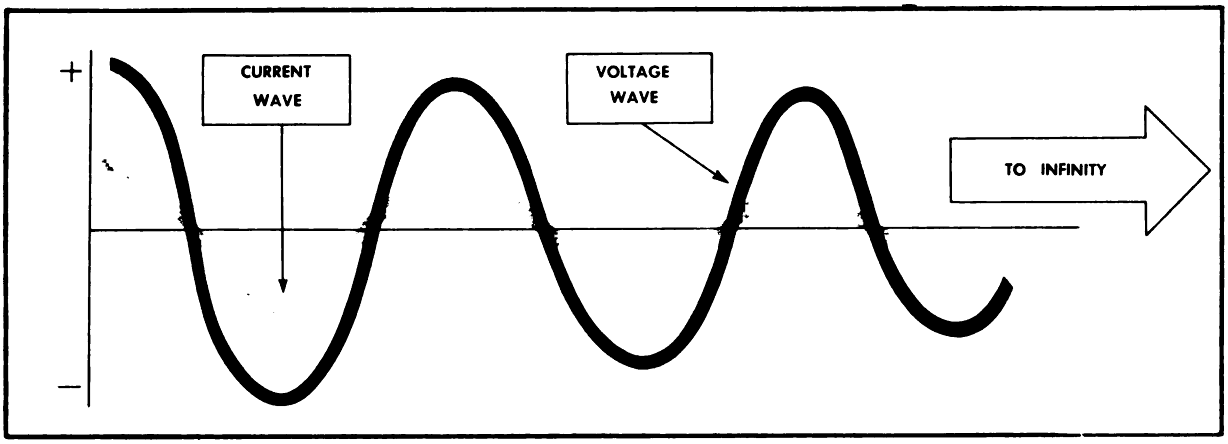
Furthermore, the impedance at any point may be resistive, reactive, or resistive and reactive, depending on the nature of the reflection. If all the energy is reflected at the output end of the transmission line, and hence no power is absorbed by the load, the line impedance is reactive at every point. If some of the wave energy is absorbed by the load and the rest is reflected, then the impedance has a resistive as well as a reactive component.

Thus you see that the impedance of a transmission line is affected by the termination of the line and the length of the line. You will understand this better after you see how the termination of a line affects the waves of voltage and current on the line.

Voltage and Current Waves on Transmission Lines

Consider what happens when a wave of RF energy is moving down a transmission line. Several different situations are considered, each dependent on the condition at the output end of the line.

TRAVELING WAVES ON AN INFINITE LINE. First consider the infinite line. This case has already been discussed earlier in this chapter. When RF energy is sent along a transmission line that is infinitely long, the ratio of voltage to current at every point remains constant along the line



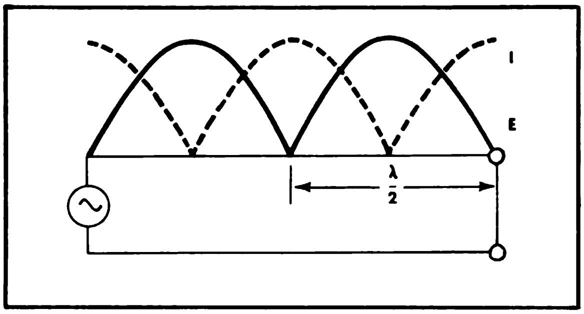
Voltage and Current Waves on an Infinite Line

and equal to the characteristic impedance. That is, the line acts simply as a resistor. The voltage and current decrease in amplitude as they travel down the line because of the losses in the line. Both are equally attenuated as a result of these losses, and these waves are in phase. Therefore, their ratio, and hence the impedance of the line, remains constant throughout the length of the line. The traveling waves of voltage and current on an infinite line may be represented as shown in the illustration above.

TRAVELING WAVES ON A PROPERLY TERMINATED LINE. A line that is terminated in an impedance or a resistance equal to its characteristic impedance behaves like an infinite line. That is, no reflections occur at the load end and the impedance is constant at every point. Since there are no reflections, all the energy (except for line losses) is absorbed by the load. Hence a line terminated in a resistive load equal to Z_0 provides maximum energy transfer. The voltage and current waves on such a transmission line are similar to those on the infinite line.

STANDING WAVES ON A SHORTED LINE. A transmission line which is shortcircuited at the output end behaves quite differently. As you know, no voltage can exist across a perfect short circuit. In order to obtain a zero voltage at the shorted end at all times, a wave whose magnitude is equal to the magnitude of the incident wave, but opposite in phase, must be reflected from the shorted end. Hence, the incident and reflected waves cancel out at the shorted end and the net voltage (or standing wave voltage) is zero at this point.

However, since both waves are traveling in opposite directions, the two voltages do not cancel out at every point on the line. At a point a quarter wavelength from the shorted end, for example, the two voltages are in phase with each other. In this case, the amplitude of the resultant standing wave is twice the amplitude of either the incident or the reflected wave alone. The illustration below, Standing Waves on a Shorted Line, shows the distribution of standing waves of voltage and current all along the line. At the point of the short circuit, the voltage is zero. The voltage rises sinusoidally to a maximum of a quarter wavelength from the short, drops to zero at the half-wavelength point, and then repeats for each half wavelength of line. Since the incident and reflected waves (which are not shown) are equal in magnitude, the standing wave voltage goes from zero to a value of twice the amplitude of the original wave. A standing wave pattern of this type indicates that all of the energy traveling down the line is reflected by the load; that is, the load does not absorb any of the energy.

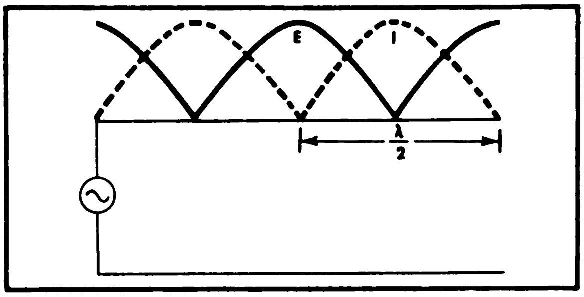


Standing Waves on a Shorted Line

The current at the short-circuited point is at a maximum value because the resistance of the short circuit is zero. Hence, the development of the standing wave of current is different from that of the standing wave of voltage. In the case of the current, the reflected wave is in phase and of equal magnitude with the incident wave at the point of the short circuit. Therefore, the standing wave of current is at its maximum amplitude at the short circuit. A quarter wavelength away from the short circuit, the current waves are 180° out of phase, and the value of the standing wave is zero. The standing wave of current, therefore, has the same form as the voltage standing wave, but is displaced a quarter wavelength (90° out of phase).

At the short circuit, where the voltage is zero and the current is at a maximum, the impedance is zero. As the ratio of voltage to current changes along the line, the impedance changes. As the voltage amplitude is rising and the current is falling, the impedance rises. At a point a quarter wavelength from the short circuit, where the voltage is maximum and the current is zero, the impedance is infinite. In the next quarter wavelength, the impedance of the line drops from infinity to zero as the voltage-current ratio is reversed. Thus, you can see that the impedance varies from point to point.

STANDING WAVES ON AN OPEN LINE. When a line is terminated in an open circuit, complete reflection of energy occurs and standing waves are developed just as in the case of the short-circuited line discussed above. However, the distribution of voltage and current is opposite to what it is for the short-circuit termination. At the open-circuited end, the voltage is at maximum amplitude and the current is zero (as shown in the illustration on this page). The voltage standing wave falls to zero a quarter wavelength away and rises to a maximum at the half-wavelength point. The current standing wave is zero at the open circuit. It rises to a maximum at the quarter wavelength point, and falls again to zero at the half wavelength point. The impedance is infinitely large at the open circuit It decreases to zero at the quarter-wavelength point, and increases to infinity again at the half wavelength point.

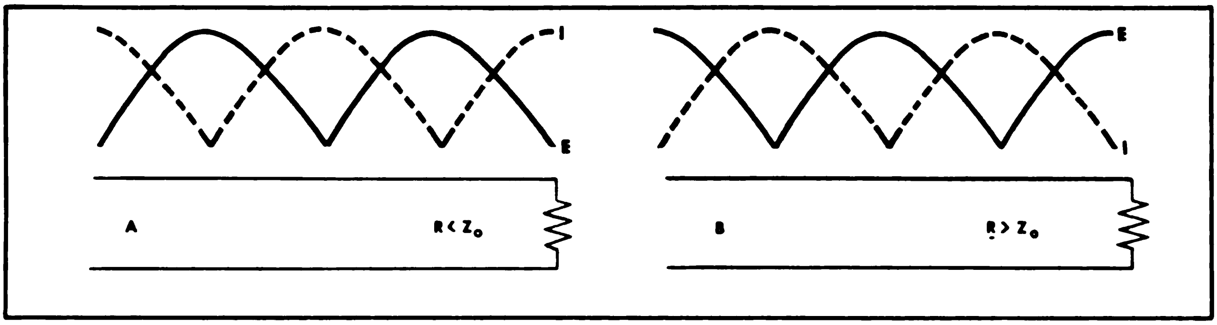


Standing Waves on an Open Line

STANDING WAVES FOR OTHER TERMINATIONS. The foregoing examples show the setting up of traveling waves on a properly terminated line, and the generation of standing waves on lines in which all the energy is reflected. There are other terminations, however, between the matched and completely mismatched conditions. In these cases, the load on the line is neither zero (short), infinite (open), nor equal to the characteristic impedance (Z_o).

Part A of the upper illustration on the next page shows the standing-wave distribution when a line is terminated in a pure resistance of a value less than Z_o. Note that the voltage and current loops and nulls are set up in the same manner as on a shorted line, but the maximum and minimum voltages and currents are different in amplitude. The amplitude does not fall to zero nor does it rise to twice the amplitude of the incident wave. Part B shows the distribution of standing waves when a transmission line is terminated in a pure resistance of a value greater than the characteristic impedance of the line. Comparing this with the representation of standing waves on an open line shows that the voltage and current loops and nulls are set up in the same manner as in an open line, but again the amplitude of the maximum and minimum voltages and currents are different.

These conditions deal with loads which are pure resistance. However, the devices which form the loads at the output ends of transmission lines usually have some reactive components. If a transmission line is terminated in a purely reactive load, standing waves will be present on the line since the load absorbs no energy. The illustration on page 365, Standing Waves on Line Terminated in a Reactance, shows these standing waves. If the line is terminated in a capacitance, as in diagram A, the current standing wave leads the voltage standing wave as



Standing Waves on Line Terminated in Resistance

seen from the output end of the line. As the capacitive load increases (that is, when the capacitance is reduced), the voltage loop approaches the output end of the line. Note that these curves are similar to those obtained on an open-circuited line, except that they do not start at the output end with the voltage maximum and the current zero.

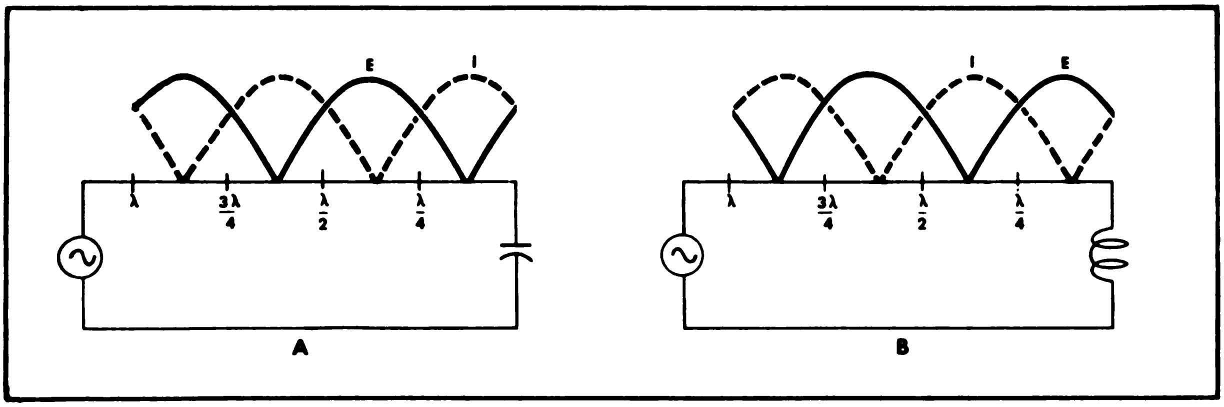
If the line is terminated in an inductance, as in diagram B, the voltage standing wave leads the current standing wave as seen from the output end. In this case, these curves are similar to those on a shorted line, except that the current at the output end does not start at maximum, and the voltage does not start at zero. As the value of inductance is increased, the voltage loop occurs nearer to the output end of the line. Therefore, if capacitive or inductive reactance is present in the load, the standing wave pattern is shifted one way or the other along the length of the line.

In practice, it is common to find that loads present both reactive and resistive components to the line. The standing waves on the line, therefore, may be considered as being a combination of the standing waves resulting from reactance and resistance alone.

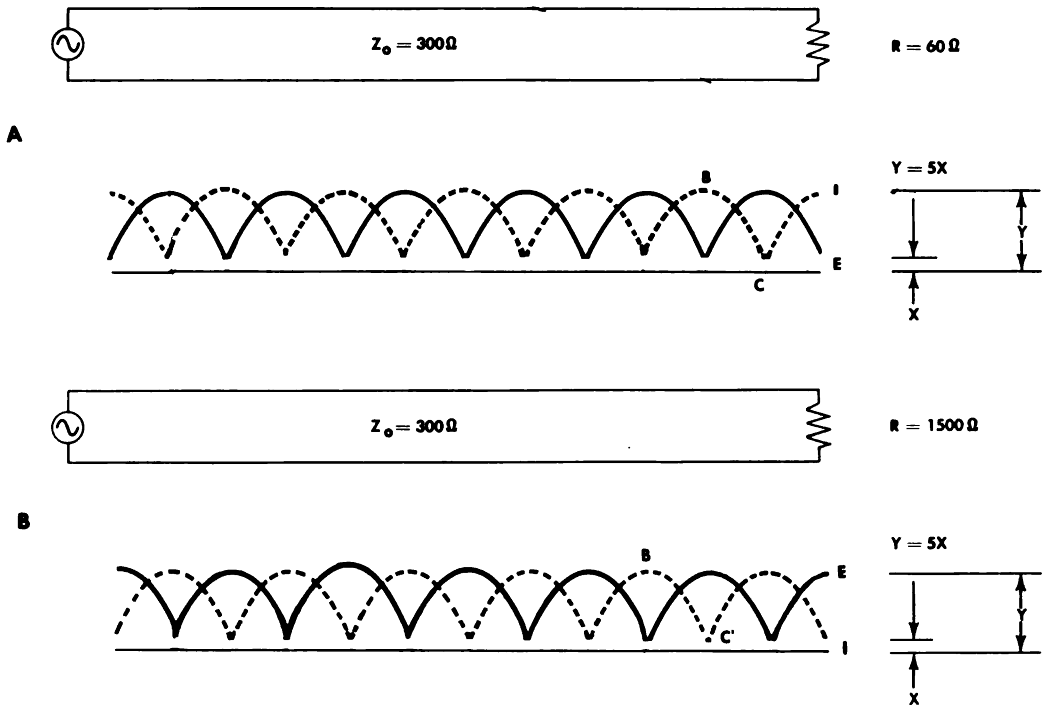
Standing-Wave Ratio

The ratio of maximum voltage to minimum voltage along a line is called the *standing-wave ratio* (SWR). The same ratio holds for maximum current and minimum current. This ratio is a measure of the amount of energy reflected or of the amount of mismatch between the load and the line. When the line is perfectly matched and all the energy is absorbed by the load, the maximum and minimum values are the same. Since there is no reflection, the current and voltage do not vary along the line. In this case, the standing-wave ratio is equal to 1.0.

Now consider the situation shown in the illustration on page 366. The Z_0 of each transmission line is 300 ohms, while the load resistance used in diagram A is 60 ohms and that used in diagram B is 1,500 ohms. If the current is measured in diagram A at points B and C and in diagram B



Standing Waves on Line Terminated in Reactance



Degree of Mismatch and SWR

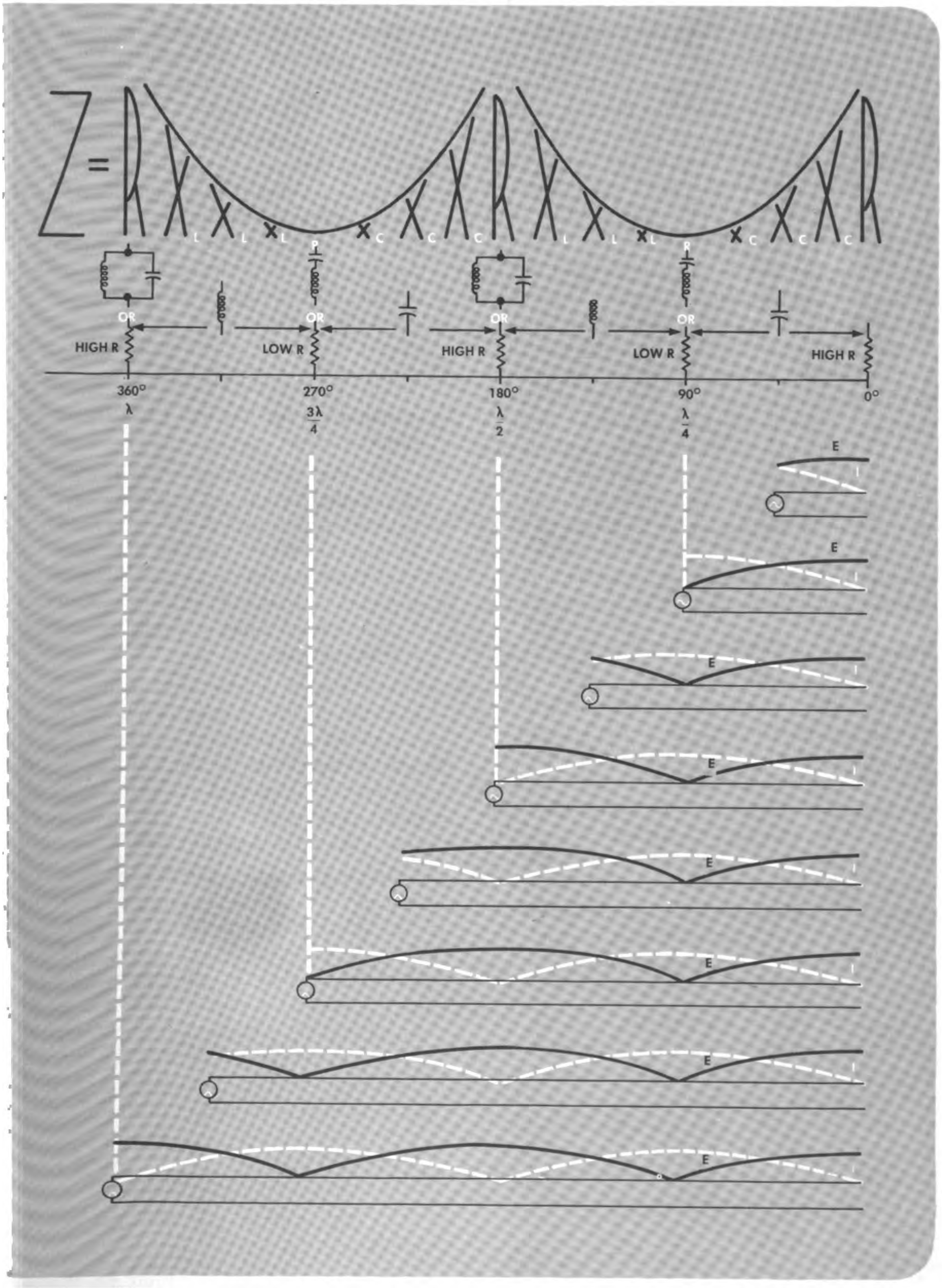
at points B′ and C′, you will find that the current at B and B′ is 5 times that measured at C and C′. In this case, the SWR is 5 to 1. This is exactly the same relationship that exists between the characteristic impedance of the line and the resistance of the load in each case.

In general, the higher the SWR, the greater the mismatch between the line and the load and the greater the power reflected in the line. In addition, a knowledge of the position of the loops and nodes along the line tells whether the load resistance is less than, or greater than, the characteristic impedance. For example, there is a voltage null (solid line) at the load in diagram A. This is so because the load resistance is less than the characteristic impedance of the line and the line behaves like a shorted line. Thus, if a standing-wave measuring device is available, it is a simple matter to determine whether the load resistance is greater or smaller than Z_0. If the load resistance is greater than Z_0, the output end of the line will appear almost like an open circuit, and a measuring device will indicate maximum voltage or minimum current at that point. If the load resistance is smaller than Z_0, the output end of the line will behave almost like a short circuit, and measurements at that point will show minimum voltage or maximum current.

RESONANT AND NONRESONANT LINES

A *nonresonant* line is a line which has no standing waves of current and voltage. Such a line is either infinitely long or terminated in its characteristic impedance. Since there are no reflections, all the energy traveling down the line is absorbed by the load. Since there are no standing waves present, this type of line is sometimes spoken of as a *flat* line. In addition, since the load impedance of such a line is equal to Z_0, no special tuning devices are required to effect a maximum power transfer and, hence, the line is also called an *untuned* line.

A *resonant* line, on the other hand, is a line which has standing waves of current and voltage. The line is of finite length and is not terminated in its characteristic impedance; hence, reflections of energy occur. The load impedance is different



Voltage, Current, and Impedance on an Open Line from the Z_0 of the line. Because of this, the input impedance may not be purely resistive but there may be reactive components present. To eliminate the reactance and to bring about maximum power transfer from the source to the line, tuning devices are used. A resonant line is therefore sometimes called a *tuned* line. Another reason for the use of this term is that the line may be used for a resonant or tuned circuit.

A resonant line is sometimes said to be resonant at an applied frequency. This means that at one frequency the line acts like a resonant circuit. It may act either as a high-resistive circuit (parallel resonant), or as a low-resistive circuit (series resonant). In order to act in this manner, the line is either open- or short-circuited at the output end and cut to some multiple of a quarter wavelength. If the length is not some multiple of a quarter wavelength, the short- or open-circuited line acts as a capacitor or as an inductor. A resonant transmission line may assume the characteristics of a resonant circuit composed of lumped capacitance and inductance.

Impedance for Various Lengths of Open Lines

Examination of the illustration on page 367 reveals the impedances for various lengths of open transmission lines. The impedance (Z) which the generator sees for various lengths of line is shown at the top. The curves above the letters of various heights indicate the relative values of the impedances presented to the generator for the various lengths. The circuit symbols indicate what the equivalent electrical circuits are for the transmission lines at each particular length. The standing waves of voltage and current are shown on each length of line.

At all odd quarter-wave points ($\frac{1}{4}\lambda$, $\frac{3}{4}\lambda$, etc.) measured from the output end of the open line, the current is maximum, the voltage is minimum, and the impedance is minimum. Thus, at all odd quarter-wave points the open-end transmission line acts like a series resonant circuit. The impedance is equivalent to a very low resistance, prevented only by small circuit losses from being zero.

At all even quarter-wave points ($\frac{1}{2}\lambda$, λ, $3/2\lambda$, etc.), the voltage is maximum, the current is minimum, and the impedance is maximum. Comparison of the transmission line with an LC resonant circuit shows that at even quarter-wavelengths, an open line acts like a parallel resonant circuit. The impedance is therefore an extremely high resistance.

In addition, resonant open lines may also act as nearly pure capacitances or inductances. The illustration shows that an open line less than a quarter-wavelength long acts as a capacitance. Also, it acts as an inductance from ¼ to ½ wavelength, as a capacitance from ½ to ¾ wavelengths, and as an inductance from ¾ to 1 wavelength, and so on. A number of open transmission lines, with their equivalent circuits, are shown in the illustration.

Impedances for Various Lengths of Shorted Lines

The shorted line may be studied with the aid of the illustration on page 369. At the odd quarter-wavelength points the voltage is high, the current is low, and the impedance is high. Since these conditions are similar to those found in a parallel resonant circuit, the shorted transmission line acts like a parallel resonant circuit at these lengths.

At the even quarter wave points, the voltage is minimum, the current is maximum, and the impedance is minimum. Since these characteristics are similar to those of a series resonant LC circuit, a shorted transmission line whose length is an even number of quarter wavelengths acts like a series resonant circuit.

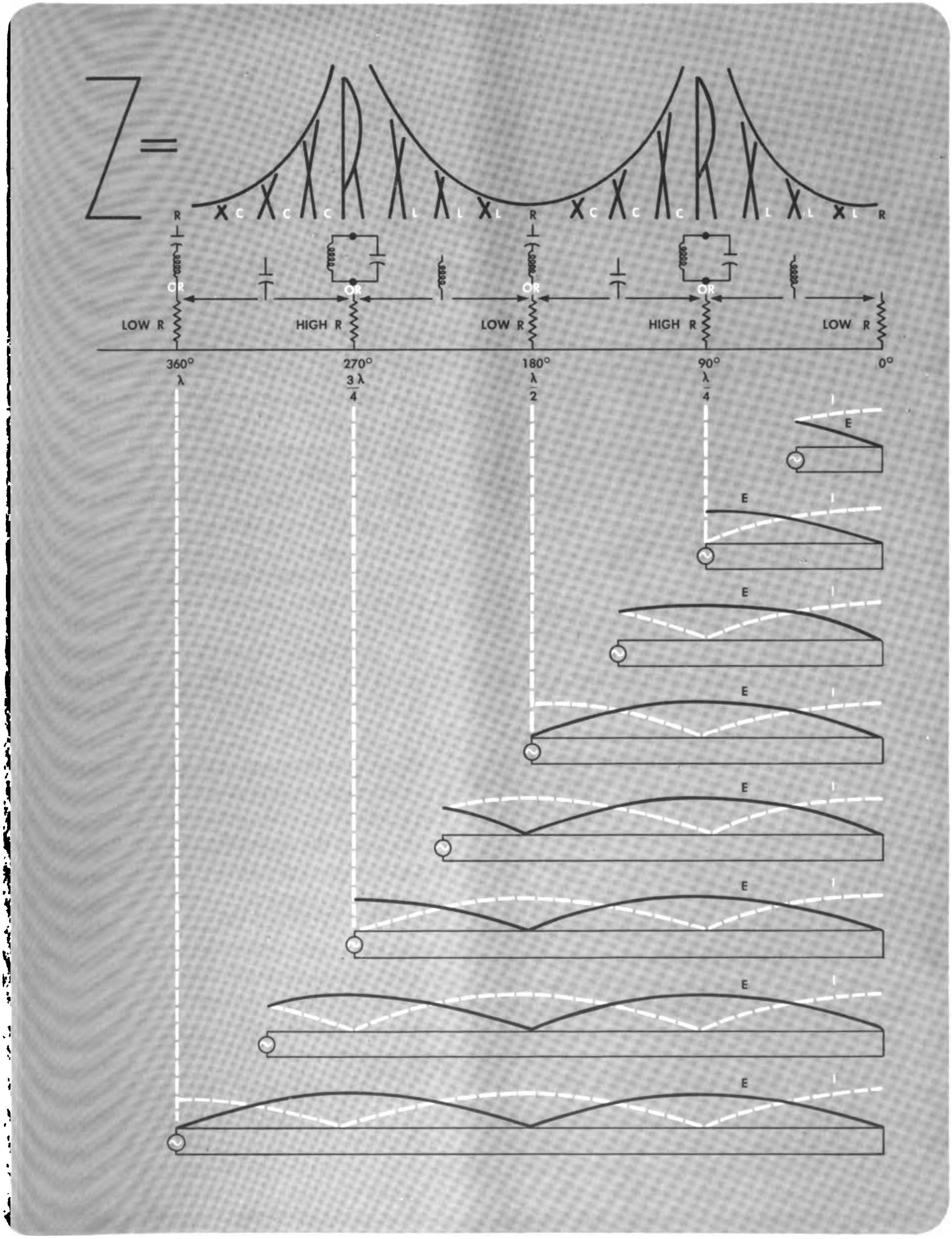
Resonant shorted lines, like the open-end lines, also may act as pure capacitances or inductances. The illustration shows that a shorted line less than ¼ wavelength acts as an inductance. A shorted line whose length is from ¼ to ½ wavelength acts as a capacitance. From ½ to ¾ wavelength it acts as an inductance, and from ¾ to 1 wavelength, as a capacitance, and so on. The equivalent circuits of shorted lines of various lengths are shown in the illustration.

Summary

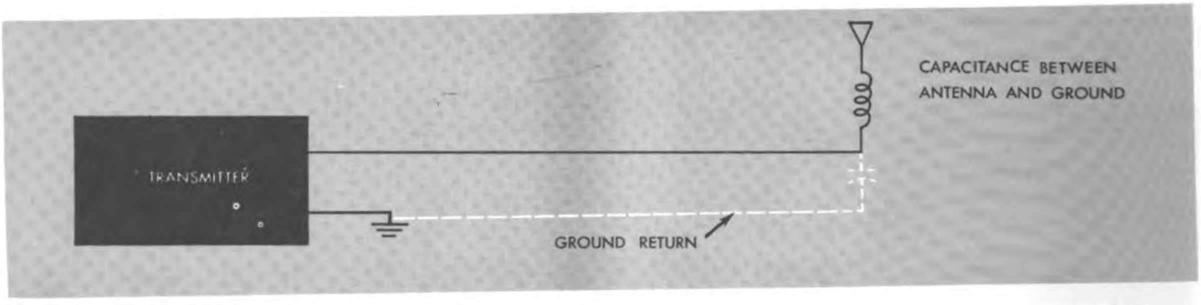
In a properly matched line:

1. The voltage and current are in phase throughout the line.

2. The ratio of the voltage to the current is constant over the entire line, and this ratio is known as the characteristic impedance.



Voltage, Current, and Impedance on Shorted Lines



Single-Wire Transmission Line

3. The input impedance and output impedance are each equal to the characteristic impedance.

4. Since the voltage and current are in phase, the line operates with minimum loss; thus, maximum power transfer is achieved.

5. The line behaves as though it were an infinite line and the standing-wave ratio is 1.0. There are no standing waves.

In an unmatched line:

1. There are standing waves present on the line.

2 The standing waves of voltage and current are 90° out of phase with each other.

3. The ratio of voltage to current varies along the line, and, thus, the impedance varies.

4. Neither the input impedance nor the output impedance is equal to the characteristic impedance.

5. Since reflections are present, the line operates with losses, and less power is transferred to the load. No power is transferred in the case of an open or shorted line.

6. Such lines may be used as parallel-resonant, series resonant, inductive, or capacitive circuits.

TYPES OF TRANSMISSION LINES

So far in this chapter, a transmission line has been illustrated as being made up of two parallel conductors. In practice, however, a transmission line may take a number of other forms. The type of line that is used depends upon the antenna used, the frequency, and equipment requirements.

Single-Wire Line

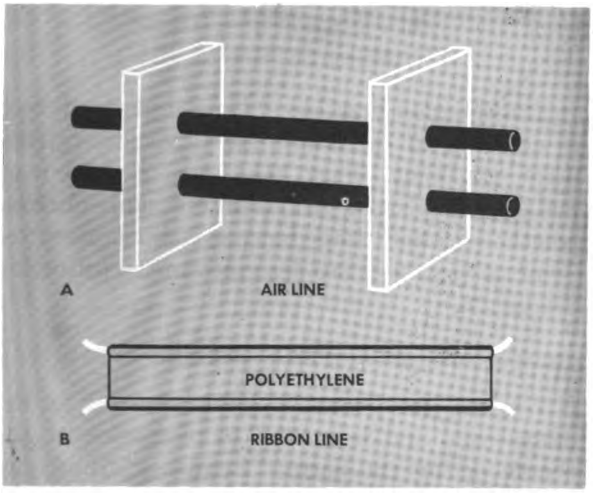
The simplest type of transmission line that can be used is the single-wire shown in the illustration above. Note that this wire uses ground, or the aircraft fuselage, as a return path. This line should be employed on a ground installation only when the site has a moist or highly conductive ground. There are two basic disadvantages of the single-wire line. One is that it is unbalanced and will therefore have a relatively large radiation loss. Second, the characteristic impedance of this line varies along its length since there is usually no fixed relationship between the wire and ground. Since there is no fixed value of Z_0, it is often difficult to match the line to an antenna.

These disadvantages are not too great when the transmission line is installed in an aircraft and a relatively low operating frequency is used. In this case, the simplicity of the single-wire line dictates its use over other types.

Two-Wire Parallel Line

The two-wire parallel line is the most common type of transmission line used for frequencies up to about 250 mc. It is relatively simple to construct and install, and it provides a reasonably good balance and constant Z_0.

This line consists of two parallel conductors which are maintained at a fixed distance from each other. To obtain a good line balance, the spacing between conductors must be constant all along the line. This is usually accomplished by insulating spacers, or spreaders, at suitable intervals as shown in diagram A, of the first illustration on page 371. In practice, such lines are



Two-Wire Transmission Lines

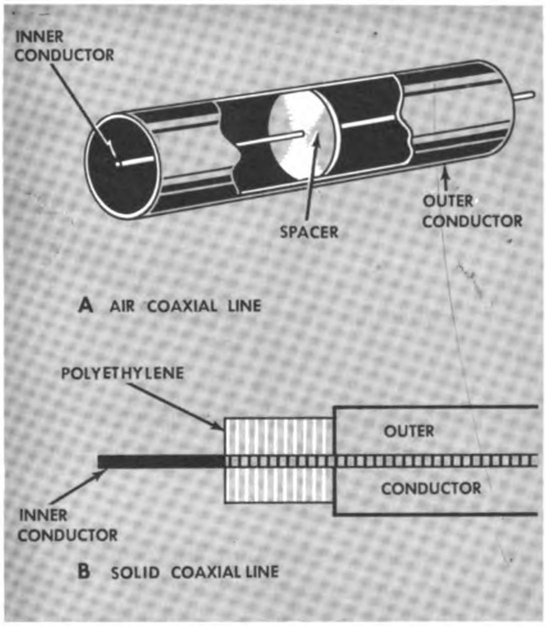
generally spaced from 2 to 6 inches apart for radio frequencies below 15 mc. The spacing for higher frequencies is usually kept to a maximum of 4 inches.

Another method of assuring a uniform spacing is to embed the conductors in a low-loss dielectric all along the length of the line, as shown in diagram B of the illustration. This type of line is often called a two-wire ribbon. It has the advantages of light weight, close and uniform spacing, flexibility, and neat appearance. Since the dielectric losses in such a line are higher than in the line which uses separating spacers with air as the dielectric, the ribbon type of line generally has a higher attenuation.

The principal disadvantage of parallel-wire transmission lines is that outside objects such as aircraft, ground, and buildings affect its balance and increase its radiation losses. This effect becomes more pronounced at higher frequencies and consequently this type of line cannot be used at these frequencies.

Coaxial (Concentric) Line

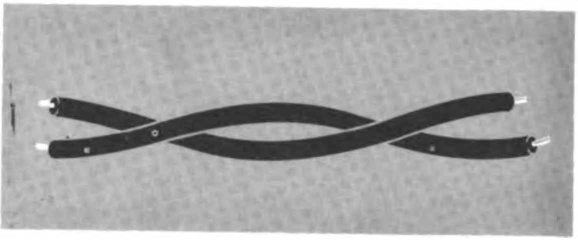
The coaxial or concentric line consists of a wire inside of, and coaxial with, a tubular outer conductor as shown in the illustration to the right above. In some cases, the inner conductor is also tubular. The inner conductor is insulated from the outer conductor by insulating spacers or beads at regular intervals. The spacers are made of Pyrex, Polystyrene, or some other material possessing good insulating qualities at high frequencies. This type of line is known as an air coaxial cable. Concentric cables are also made with the inner conductor of flexible wire which is insulated from the outer conductor by a solid and continuous insulating material. This type of line is called a solid coaxial cable. Flexibility of the cable may be gained if the outer conductor is made of metal braid. The losses in this latter type of coaxial line are rather high.



Construction of a Coaxial Cable

The chief value of the coaxial line is that it minimizes radiation losses. In the two-wire parallel line, the electric and magnetic fields can extend out into space for great distances. In coaxial construction, however, no electric or magnetic fields extend outside of the outer conductor and, being confined, all fields exist within the space between the two conductors. Thus, the coaxial line is a shielded line.

The advantage of an efficient concentric line is that it has practically no radiation loss and, therefore, nearby objects cause no difficulty in operation. Some of its disadvantages are: it is more expensive for a given length of line, its attenuation at extremely high frequencies limits its use except for short distances, and the interior must be kept dry in order to prevent excessive leakage between the conductors. To prevent



Twisted Pair Transmission Line

condensation of moisture, the air concentric lines may be filled with dry nitrogen at low pressures.

Twisted Pair

A twisted-pair transmission line consists of two insulated wires twisted together as shown in the illustration above. This forms a flexible line without the use of spacers. A twisted-pair line should always be used as a nonresonant line, because standing waves on the line will produce excessive losses and can easily break down the insulation between conductors at voltage loops.

Such a line is often used as a short link between stages in a transmitter. Above about 15 mc, however, the insulation causes appreciable dielectric losses and the twisted-pair line should not be used.

Shielded Pair

The shielded pair consists of two parallel conductors separated from each other and surrounded by an insulating dielectric material, as shown in the illustration to the right. The conductors are enclosed in a copper-braid tubing which acts as a shield. The outer layer consists of a rubber or flexible composition coating to protect the line against moisture and friction. The advantages of the shielded pair are that the two conductors are balanced to ground and that the wires are shielded against radiation.

Waveguides

At frequencies above about 1,000 mc, a different type of transmission line, known as a waveguide, is used. Waveguides are discussed more fully in chapter 36.

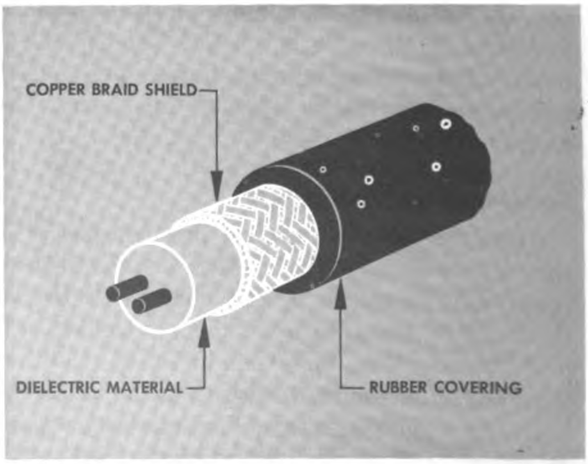
APPLICATIONS OF TRANSMISSION LINES

Transmission lines have many important uses in addition to the transmission of power from point to point. They may also be used as metallic insulators, wave filters and chokes, reactances, and impedance-matching devices. Some of these uses are described in the following discussion.

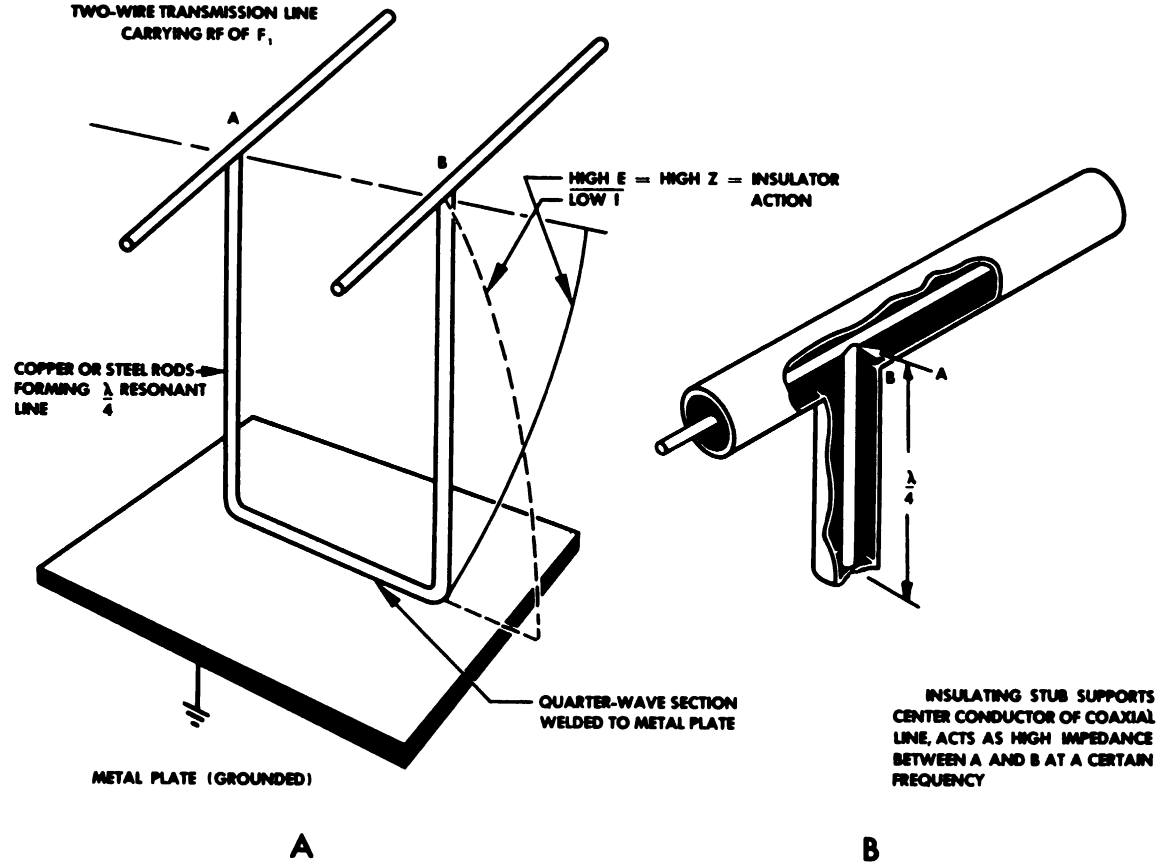
Quarter-Wave Line as a Metallic Insulator

When a quarter-wave line is shorted at the output end and is excited to resonance at the other end by the correct frequency, there are standing waves of current and voltage on the line. At the short circuit, the voltage is zero, while the current is maximum. At the input end, the current is nearly zero and the voltage is maximum. Therefore, the input impedance is very great. An exceedingly high impedance across the terminals looks like an insulator to another line connected to the input end. Therefore, such a line can be used as an insulator at its two open terminals. Note that it will act as such an insulator at one frequency only.

The illustration on the next page shows a quarter-wave section of line acting as a standoff insulator for a two-wire transmission line. Naturally, for direct current, this section acts as a short circuit across the line; but at the particular frequency for which this section represents a quarter wave-length, it acts as a highly efficient insulator. At terminals A and B, there is a high voltage and a low current. Since $Z=E/I$, the



Shielded Pair Transmission Line


TWO-WIRE TRANSMISSION LINE CARRYING RF OF F₁
A
B
HIGH E / LOW I = HIGH Z = INSULATOR ACTION
COPPER OR STEEL RODS FORMING λ/4 RESONANT LINE
QUARTER-WAVE SECTION WELDED TO METAL PLATE
METAL PLATE (GROUNDED)
B
A
λ/4
INSULATING STUB SUPPORTS CENTER CONDUCTOR OF COAXIAL LINE, ACTS AS HIGH IMPEDANCE BETWEEN A AND B AT A CERTAIN FREQUENCY
A
B

Quarter-Wave Transformer

impedance between A and B must be very high, just as though a conventional insulator were used. If the frequency of the signal transmitted down the line varies greatly, the section becomes a poor insulator across the line.

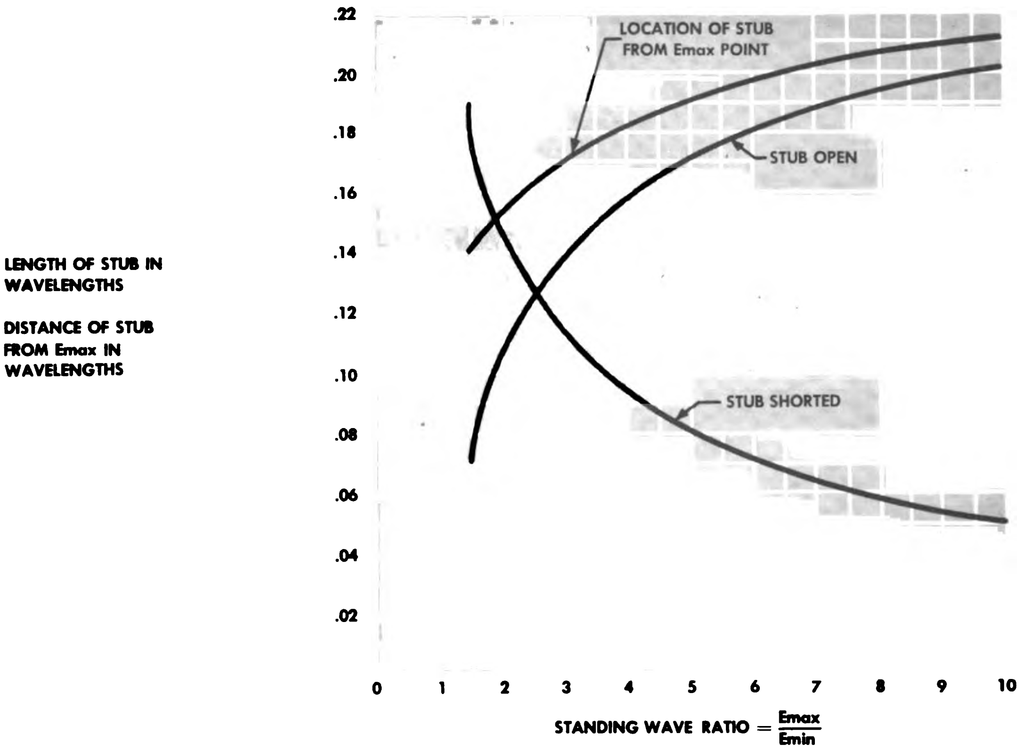
An important application of the quarter-wave insulator is the "teeing" of coaxial lines. You have noted that the inner conductor of a coaxial line must be separated and insulated from the outer conductor. This is generally done by either filling the space between the two conductors with a dielectric material, or using dielectric beads or spacers. However, the dielectric insulator has the disadvantage of introducing dielectric losses in the line, especially at the higher frequencies. The use of a quarter-wave stub to separate and insulate the inner and outer conductors, as shown in part B, reduces this dielectric loss. Consequently, if the quarter-wave stubs, or "tees," as they called, are spaced close enough to provide adequate mechanical support, this type of insulation is usually more efficient than spacers, provided that the line is operated at only one frequency or over a narrow band of frequencies.

Stub Matching

It is possible to match a line to an antenna by first determining the point along the line where the resistive component is equal to Z_0 and then tuning out the reactance that exists at this point. The reactance can be tuned out by using a section of line, or ***stub***. If the line reactance is inductive, a capacitive stub must be used. If the line is capacitive, an inductive stub must be used.

The problem is to determine exactly where along the line the resistive component is equal to Z_0 and what length of stub is required. This can be done with the aid of a standing-wave meter and the chart on the next page illustrated. The procedure for determining the position and length of the stub required is best understood by working out a typical problem.

With a standing-wave meter, determine the standing waves that exist along the line as shown in diagram A of the illustration on page 375. Compute the standing-wave ratio by dividing the maximum voltage (E_{max}) by the minimum voltage (E_{min}). Since E_{max} is 50 volts and E_{min} is



Impedance Matching with a Stub

10 volts, the standing wave ratio is 5. Examine the vertical line on the chart, Impedance Matching with a Stub, corresponding to a standing-wave ratio of 5. This line crosses three curves. One curve indicates the length of the short-circuited stub required, the middle curve indicates the length of the open-circuited stub required, and the top curve indicates the position of the stub with respect to the voltage loop. In this case, the vertical line crosses the stub location curve at 0.19 wavelength. This means that the stub, either open or shorted, must be located 0.19 wavelength from the voltage loop nearest the antenna as shown in diagram B of the illustration. If the stub is short-circuited, it must be located at a point 0.19 wavelength toward the input end of the line. If the stub is open circuited, it must be located 0.19 wavelength toward the output end of the line.

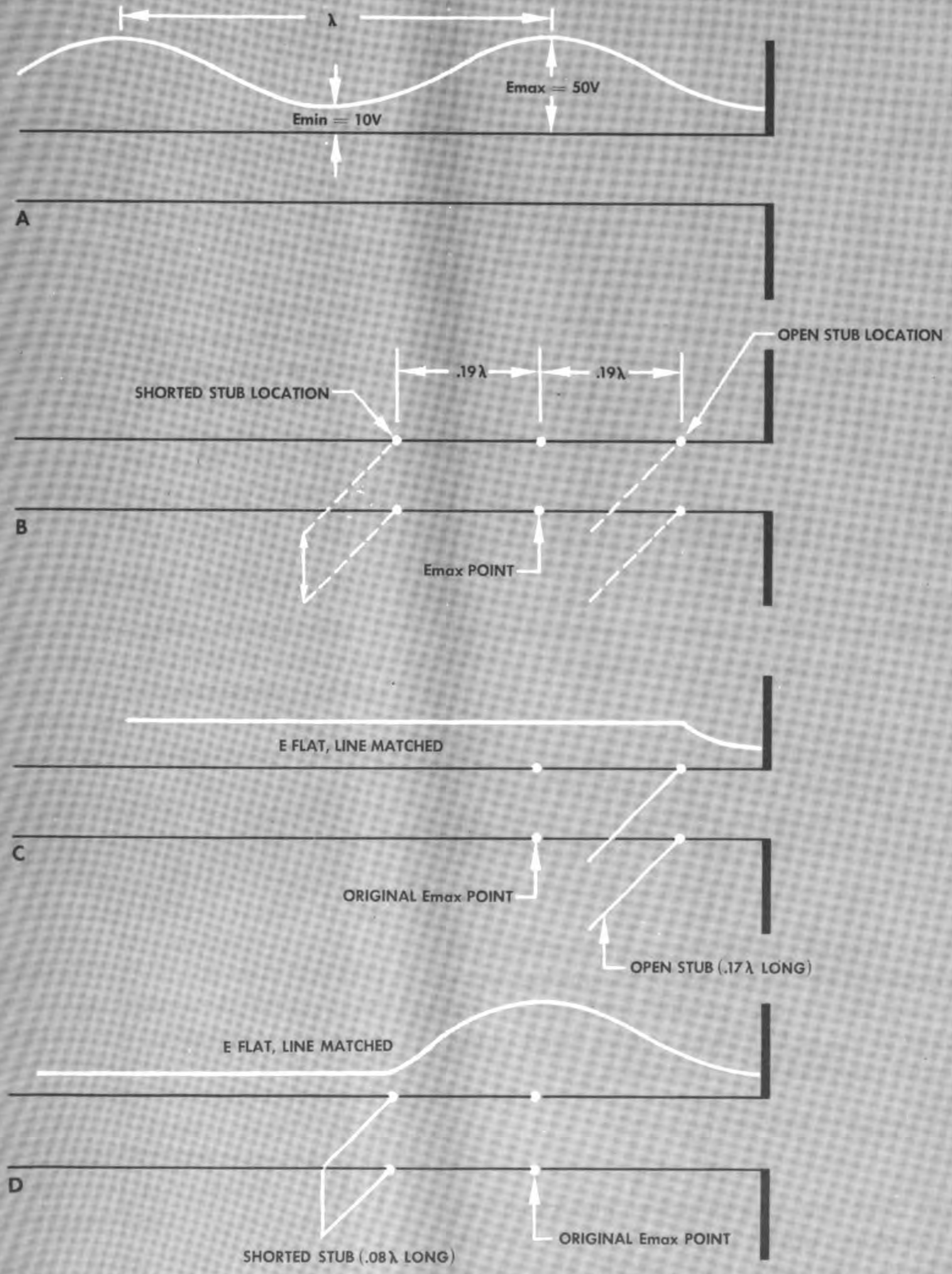
The length of the open-circuit stub required is determined by the open stub curve in the chart previously given. In this case the stub must be 0.17 wavelength long. If such a stub is placed 0.19 wavelength from the voltage maximum toward the antenna, the line will be matched as shown in diagram C.

A shorted stub can be used instead of the open stub. In this case the stub length, as obtained from the chart, is 0.08 wavelength. This stub is placed 0.19 wavelength toward the line input as shown in diagram D. In general, the shortest possible stub is used. Hence, in this illustration, it would be desirable to use the shorted stub, rather than the open stub.

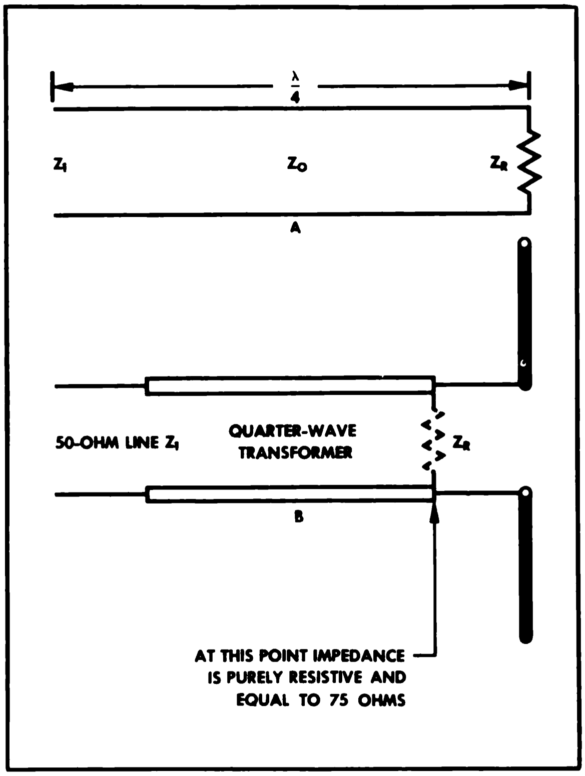
In a similar manner, the stub position and length for any measured standing-wave ratio can be determined. Furthermore, the principles of stub matching are applicable to any type of transmission line, including the coaxial line and the waveguide.

Impedance Transformers

A second method of matching the transmission line to the antenna is to use an impedance transformer. In this case, a point is located along the line where the reactance is zero and the line has only a resistive component. An impedance transformer is used to match this resistive component to the characteristic impedance of the line. A transmission-line element often used for this purpose is a device called a quarter-wave transformer.



Stub Matching Problem


λ/4
Z1
Z0
ZR
A
50-OHM LINE Z1
QUARTER-WAVE TRANSFORMER
ZR
B
AT THIS POINT IMPEDANCE IS PURELY RESISTIVE AND EQUAL TO 75 OHMS

Quarter-Wave Transformer

If a quarter-wave section of line has an impedance Z_R placed across its output as shown in A of the illustration above, the input impedance to this line may be found from the

$$Z_1 = \frac{Z_0^2}{Z_R}$$

formula, where Z_1 is the input impedance, Z_0 is the characteristic impedance of the quarter-wave line, and Z_R is the output impedance. For example if Z_R is 100 ohms and Z_0 is 50 ohms, the input impedance to this line is

$$Z_1 = \frac{(50)^2}{100}$$
$$= 25 \text{ ohms}$$

Note that the quarter-wave line transforms the 100-ohm output impedance to a 25-ohm impedance. In many applications, the output impedance and the input impedance are known, and it is desired to determine what value of Z_0 is required to effect the transformation. For these applications, you can express the equation in the following manner,

$$Z_0 = \sqrt{Z_R \times Z_1}$$

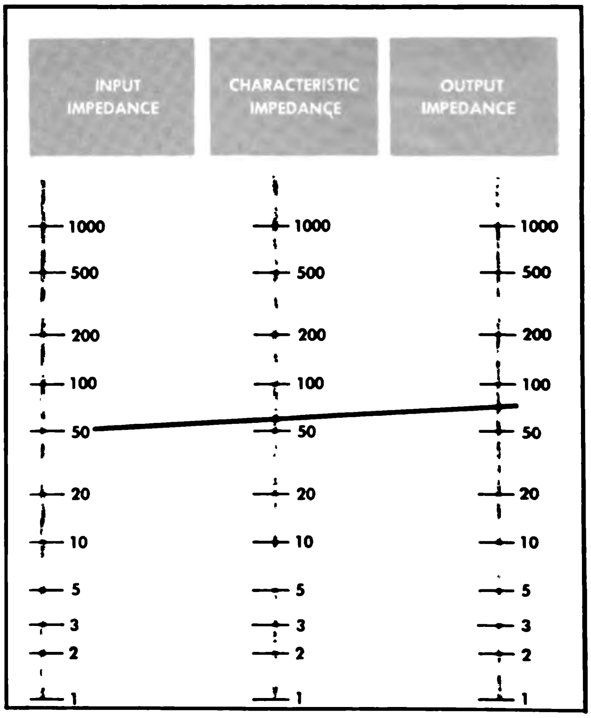
For example, in diagram B of the illustration to the left it is necessary to match a 50-ohm line to a 75-ohm load impedance. A quarter-wave transformer is required which is constructed of a transmission line whose characteristic impedance is

$$Z_0 = \sqrt{50 \times 75}$$
$$= \text{about } 61 \text{ ohms}$$

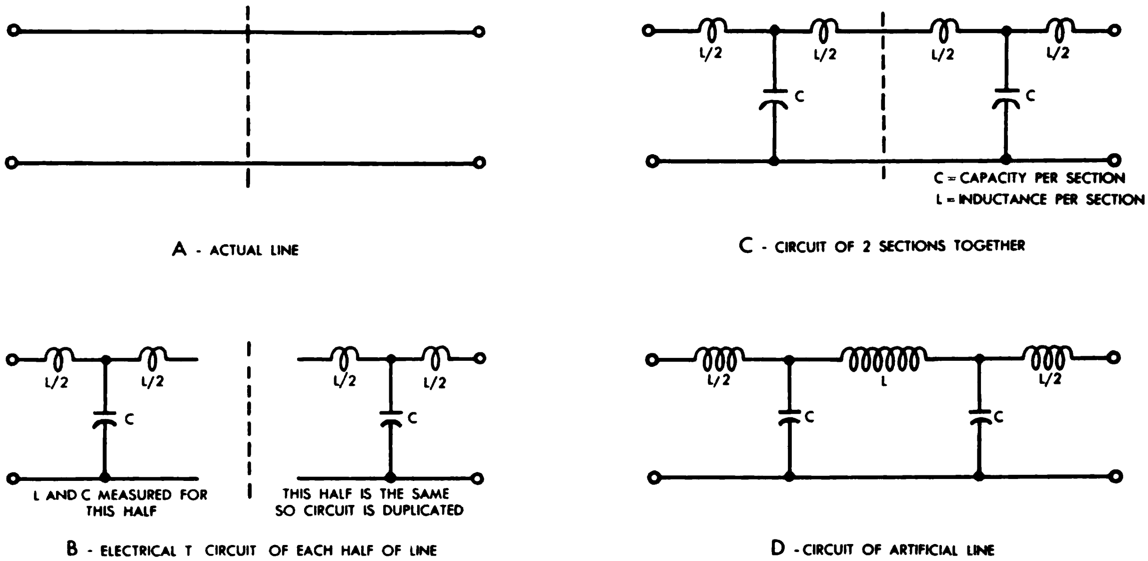
The nomograph shown below enables you to obtain Z_0 for a quarter-wave transformer by means of a straightedge. To use this chart, place a straightedge on the values of Z_1 and Z_R that are used. The point at which this straightedge crosses the center line corresponds to the required characteristic impedance. Consider the previous example, where Z_1 is 50 ohms and Z_R is 75 ohms. A line is drawn between these two points on the chart. This line intersects the Z_0 line at about 60 ohms. Therefore, the quarter-wave transformer characteristic impedance must be this value, as was previously calculated.

ARTIFICIAL TRANSMISSION LINES

Although the delay characteristics and charging action of RF transmission lines make them useful in some radar applications, they would be too long for other applications. An actual RF



Quarter-Wave Transformer Nomograph


L/2
L/2
C
C = CAPACITY PER SECTION
L = INDUCTANCE PER SECTION
L AND C MEASURED FOR THIS HALF
THIS HALF IS THE SAME SO CIRCUIT IS DUPLICATED
L

A - ACTUAL LINE

C - CIRCUIT OF 2 SECTIONS TOGETHER

B - ELECTRICAL T CIRCUIT OF EACH HALF OF LINE

D - CIRCUIT OF ARTIFICIAL LINE

Design of Artificial Transmission Line

line would be approximately several hundred feet long for some uses in radar. Even if a line of such length could be wound in a coil, it would occupy more space than is available in an aircraft. For this reason artificial lines which possess all the characteristics of actual lines without the bulk and physical length of an actual line are used. Artificial lines are constructed by first determining the capacity and inductance needed for the desired line and then constructing the line with equivalent lumped inductors and capacitors. The result is an electrically equivalent, but physically different line.

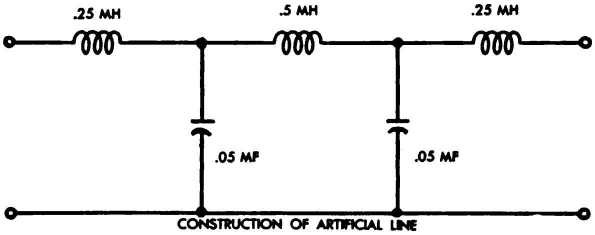
At A of the illustration above is shown an actual two-wire line in two sections. At B is shown the equivalent circuit for each section with the characteristics measured. Since both sections have identical characteristics, the two can be joined together as shown at C. The final circuit, shown at D, lumps the inductive and capacitive characteristics of each section and forms the artificial line. As you can see, the artificial line performs just as an actual line, but does not occupy as much space.

Introducing Delay

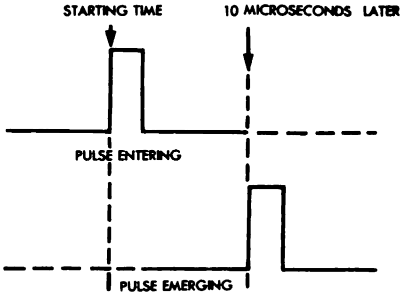
In any line, time is required for any voltage change to travel the length of a line. This time characteristic makes it possible to delay the transfer of a voltage change in its travel from one circuit to another. The artificial line shown in the illustration at the right provides a means of delaying a pulse voltage 10 microseconds. This line is constructed with an inductance of .5 millihenry and a capacity of 0.05 mf in each of the two sections. To compute the time for a voltage change to move through an artificial line, use the following equation:

$$T_d = N\sqrt{LC}$$

where T_d is the time delay in seconds for the entire line, N is the number of sections, L the inductance per section in henries, and C the capacitance per section in farads.



CONSTRUCTION OF ARTIFICIAL LINE



EFFECT OF LINE ON VOLTAGE CHANGE

Delaying a Pulse 10 Microseconds with Artificial Line

The following gives the calculations for finding the time delay for the artificial line:

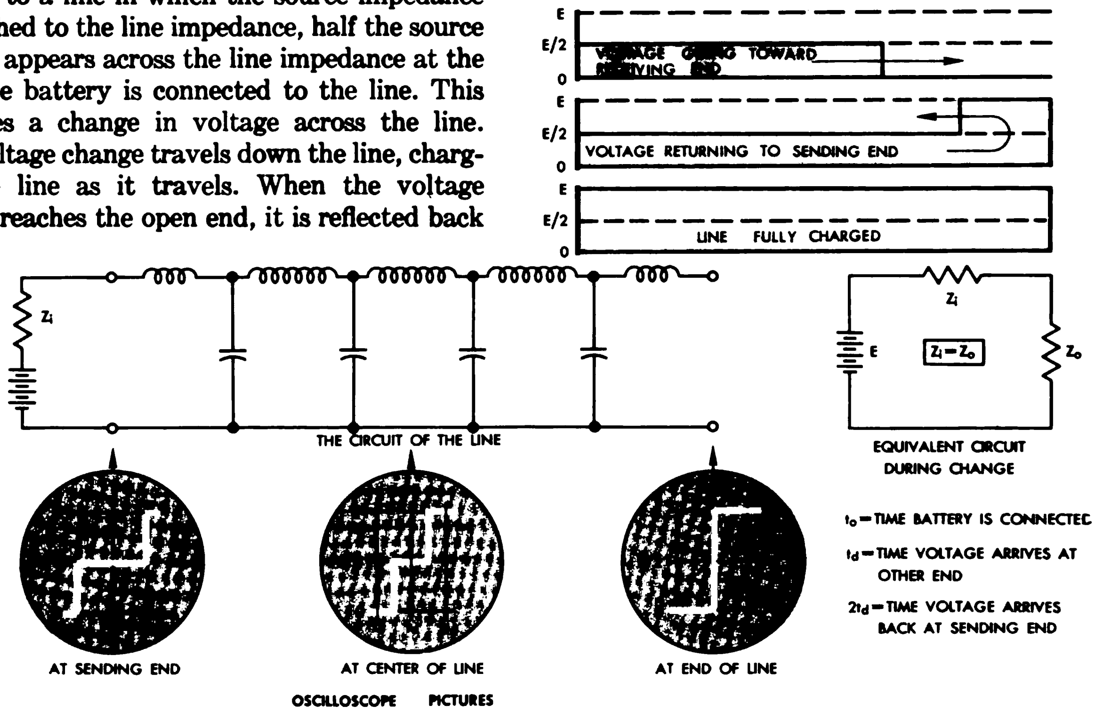
$$
\begin{aligned}
T_d = N\sqrt{LC} &= 2\times\sqrt{.5\times10^{-3}\times.05\times10^{-6}}\\
&= 2\times\sqrt{5\times10^{-4}\times5\times10^{-8}}\\
&= 2\times\sqrt{25\times10^{-12}}\\
&= 2\times5\times10^{-6}\\
&= 10\times10^{-6}\text{ seconds}\\
&= 10\text{ microseconds}
\end{aligned}
$$

One radar set uses the time delay characteristic of an RF line in its operation. In this set, the time base is generated and does not start until after the first radar echo returns to the receiver. In order for these signals to be displayed at their correct range, they are sent through an artificial RF delay line. This line delays the signals just enough so that the very first returning signal does not arrive at the cathode-ray tube (CRT) until the time base begins. Although a large number of video frequencies are present in the signal going through the line, all of them are delayed an equal amount.

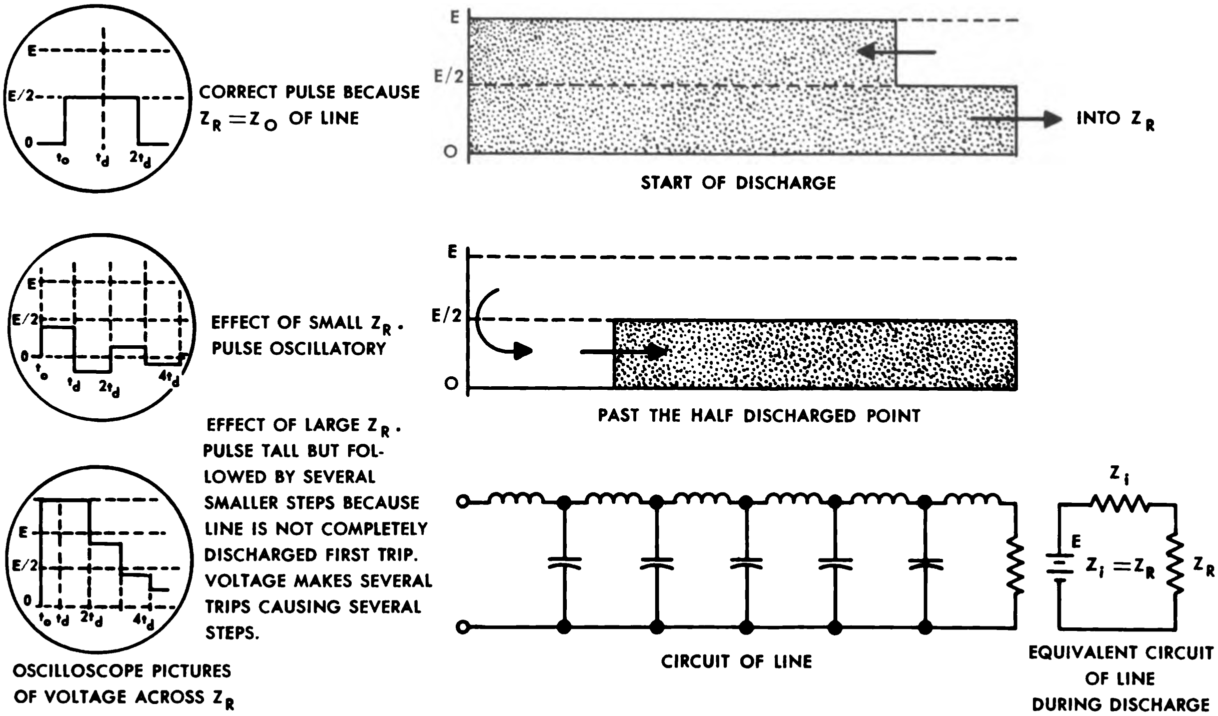
Charging and Discharging a Line

Another phenomenon of RF lines is the manner in which the lines charge and discharge when a DC voltage is applied to them. When DC is applied to a line in which the source impedance is matched to the line impedance, half the source voltage appears across the line impedance at the time the battery is connected to the line. This produces a change in voltage across the line. This voltage change travels down the line, charging the line as it travels. When the voltage change reaches the open end, it is reflected back along the line. All capacitors charge to half the battery voltage when the voltage travels down the line and to full battery voltage when the voltage is coming back. When all capacitors have been charged to a value equal to the battery voltage, current stops flowing from the battery. If you view these voltage changes with an oscilloscope, their waveshapes will appear as shown in the illustration below. Note that a voltage which is shorter in time nearer the receiving end than at the sending end is produced. This voltage is called a step voltage.

When this line is discharged into a resistance equal to its characteristic impedance, the output is a square wave which has a constant amplitude and a duration equal to twice the time for the pulse to travel the length of the line. The amplitude is equal to half the charge voltage because the line discharges through its own impedance and the load impedance (Z_R), which are connected in series. This discharge produces a drop of voltage which immediately travels down the line and back to the starting place as shown in the first illustration on page 379. As the end of the line is open, the reflection is in phase and dis-


E
E/2
0
VOLTAGE GOING TOWARD RECEIVING END
E
E/2
0
VOLTAGE RETURNING TO SENDING END
E
E/2
0
LINE FULLY CHARGED
Zi
THE CIRCUIT OF THE LINE
AT SENDING END
AT CENTER OF LINE
AT END OF LINE
OSCILLOSCOPE PICTURES
Zi
E
Zi=Zo
Zo
EQUIVALENT CIRCUIT DURING CHARGE
t0—TIME BATTERY IS CONNECTED
td—TIME VOLTAGE ARRIVES AT OTHER END
2td—TIME VOLTAGE ARRIVES BACK AT SENDING END

Charging Artificial Transmission Line


E
E/2
0
t_o
t_d
$2t_d$
CORRECT PULSE BECAUSE $Z_R = Z_O$ OF LINE
E
E/2
O
INTO Z_R
START OF DISCHARGE
$4t_d$
EFFECT OF SMALL Z_R. PULSE OSCILLATORY
PAST THE HALF DISCHARGED POINT
EFFECT OF LARGE Z_R. PULSE TALL BUT FOLLOWED BY SEVERAL SMALLER STEPS BECAUSE LINE IS NOT COMPLETELY DISCHARGED FIRST TRIP. VOLTAGE MAKES SEVERAL TRIPS CAUSING SEVERAL STEPS.
OSCILLOSCOPE PICTURES OF VOLTAGE ACROSS Z_R
CIRCUIT OF LINE
Z_i
$Z_i = Z_R$
Z_R
EQUIVALENT CIRCUIT OF LINE DURING DISCHARGE

Discharging an Artificial Transmission Line through Z_R

charges the line as it travels back. The output voltage across the load impedance (Z_R) is equal to half the charge voltage and lasts twice the time of travel on the line.

Usually, an artificial line is constructed to generate a square wave. When L and C for each section are known, the time duration of the output square pulse, T_p, is

$$T_p = 2N\sqrt{LC}$$

The duration of the output pulse is equal to twice the delay that can be introduced with an artificial line. The time of the output pulse that can be obtained with the line on page 377 is, therefore, 20 microseconds.

When an artificial line is not terminated in its characteristic impedance (Z_o), several trips (reflections) are required for complete discharge. This means that discharge cannot be completed in a single round trip. In the illustration above, three oscilloscope pictures show the discharge action of an artificial line across different values of Z_R.

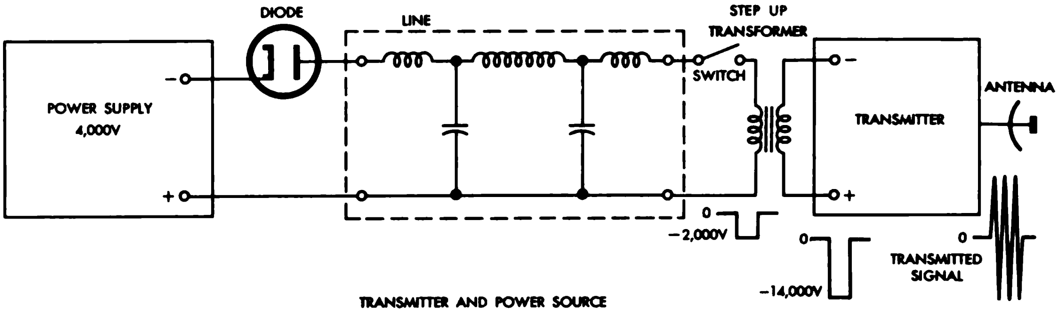
When Z_R is small, the change that is reflected not only cancels the voltage left on the line, but also recharges the line to a smaller voltage in the opposite direction, producing an oscillatory output waveshape. When Z_R is large, the amplitude of the pulse is large, but it is followed by several smaller steps called tails. These tails exist because the line does not discharge completely during the first trip and more trips are necessary for complete discharge. Note in the upper diagram the correct pulse obtained when Z_R equals the Z_o of the line.

To calculate Z_o for an artificial line, use the formula

$$Z_o = \sqrt{\frac{L}{C}}$$

The following are the calculations for finding Z_o for the artificial line of two sections shown on page 377:

$$\begin{aligned} Z_o &= \sqrt{\frac{L}{C}} = \sqrt{\frac{.5 \times 10^{-3}}{5 \times 10^{-8}}} \\ &= \sqrt{\frac{5 \times 10^{-4}}{5 \times 10^{-8}}} \\ &= \sqrt{10^4} \\ &= 100 \text{ ohms} \end{aligned}$$

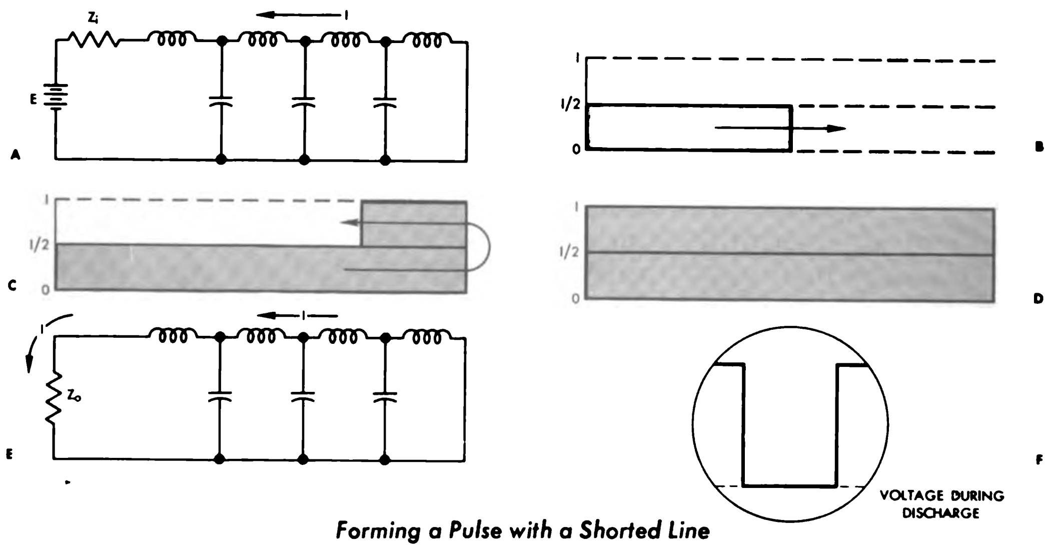


Block Diagram Showing Pulse Line Used in Radar Transmitter Circuits

Use of Artificial Line

An artificial transmission line is capable of producing square pulses several thousand volts in amplitude. The illustration above is a simplified drawing of such a pulse line in a radar transmitter. The 4000-volt power supply charges the line; the series diode prevents the line from discharging through the power supply. When the switch is closed, a 1-microsecond, 2000-volt pulse is generated across the primary of the step-up transformer. The 14,000-volt pulse developed across the secondary serves as cathode voltage for the transmitter. During the time the cathode voltage is applied (one microsecond), a 1-microsecond pulse of RF energy is produced.

A short-circuited artificial line can also be used to form pulses. In this case, operation is based upon the energy that is stored in the magnetic fields around the coils. Notice in diagram A below that when the battery voltage E is applied to the line, current flows in the line, charging the capacitor. At the short-circuited end, the current is reflected in phase and returns up the line (diagrams B and C). The capacitors discharge through the coils; thus the value of the current on the return trip is double the current that went down the line (diagram D). Although the voltage across the capacitors is zero, the current through the coils is maximum. If at the instant the line is fully charged, a load (Z_0) equal to the line impedance replaces the battery and its internal impedance (Z_i), the current due to the collapsing field about the coils flows through the load. The voltage output pulse across Z_0 is shown in diagram F.



Forming a Pulse with a Shorted Line

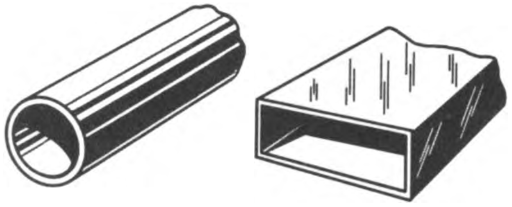
Chapter 36

Waveguides, Cavity Resonators and Duplexers

Some of the electronic equipment used by the Air Force operates at frequencies above 1,000 mc. Because of the very short wavelengths of these frequencies, they are called *microwaves*. Transmission lines employed at these frequencies are considerably different physically and electrically from those previously discussed.

The reason for using special transmission lines at microwave frequencies is that the tendency for conductors to radiate the energy they carry increases as the frequency increases. This radiation effect increases to a point where, if a microwave signal were applied to a two-wire line, for example, most of the energy would be radiated into space and practically none would reach the output end.

You can use a coaxial line for microwave transmission because, even though the inner conductor radiates energy, all of this energy is kept within the confines of the outer and inner conductors; it cannot escape into space. A much simpler arrangement, and one which has a number of advantages is to take out the center conductor and transmit the wave through the remaining conductor. When this is done, the resulting transmission line is called a cylindrical waveguide. This type of transmission line is shown at A of the illustration on the following page. When a hollow rectangular conductor, as shown at B, is used instead of a hollow round one, the transmission line is called a rectangular waveguide.



Cylindrical and Rectangular Waveguides

In energizing a waveguide, an antenna is used to produce the required electromagnetic waves in exactly the same way that an antenna is used to radiate energy into free space. (Antennas and radiation are discussed in Chapter 37.) In order to pick up the energy from the waveguide, a receiving antenna is located at the output end of the guide to pick up the electromagnetic waves in the same way as in free space. The use of an input antenna to energize the wave guide and of an output antenna to pick up energy from the waveguide is illustrated to the right below.

The use of an antenna in a waveguide, as contrasted with the use of an antenna radiating in space, makes it possible to confine and to guide the electromagnetic energy along a definite path. Energy radiated into free space is not usually confined to a definite path; hence, such waves are called unguided waves. Those traveling in waveguides, on the other hand, are called guided waves.

WAVEGUIDES

Waveguides have certain advantages and disadvantages when used to transmit microwaves. The first advantage to be discussed, the fact that there is no radiation loss, indicates why waveguides are preferred to open wire transmission lines. The last four advantages show why they are preferred to coaxial lines. The disadvantages of waveguides are also pointed out; from these you will see why they are used only for microwaves.

Advantages

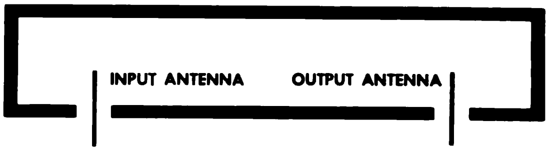
A waveguide has no radiation loss. At microwave frequencies ordinary conductors radiate most of the energy applied to them. Consequently in the case of two-wire or single-wire transmission lines, large radiation losses occur. In waveguides, however, radiation losses are practically zero because all of the electromagnetic energy is confined inside the guide. Radiation losses in coaxial lines are also very low, and for this reason, coaxial lines are sometimes used for microwave transmission.

A waveguide has no dielectric loss. One of the major reasons for signal attenuation in a coaxial line is the leakage in the dielectric used to support the inner conductor. However, because the dielectric material has some resistance, some of the energy transmitted down the line is absorbed by the insulator.

The frequency of the transmitted energy is the principal factor governing dielectric loss. As the frequency is increased, the amount of electromagnetic energy absorbed by the dielectric also increases. The dielectric loss is therefore much higher at microwave frequencies than at the lower radio frequencies. Dielectric loss of energy is eliminated in a waveguide, since there is no center conductor requiring a solid dielectric support.

A waveguide has low skin-effect loss. Whenever an electromagnetic wave strikes a conductor, a current is induced in that conductor, and, because every conductor has some resistance, energy is dissipated in the form of heat. The amount of energy lost by electromagnetic waves striking the conductor is directly proportional to the effective resistance of the conductor. If the resistance is high, the losses in the conductor are also high. The effective resistance of a conductor is determined by its DC resistance, the cross-sectional area, and the frequency of the energy being transmitted.

In a coaxial cable, the DC resistance of the inner conductor is very small. The cross-sectional area of the inner conductor, however, is smaller than the cross-sectional area of the outer conductor, and, since the resistance of a conductor



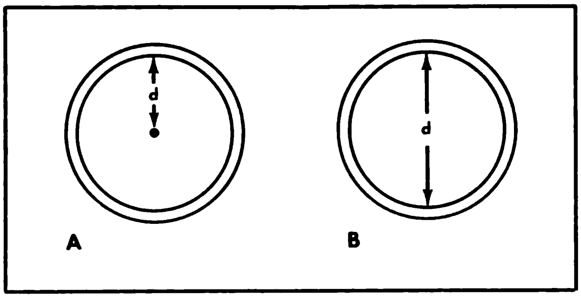
Energizing and Picking up Energy from Rectangular Waveguide is inversely proportional to the cross-sectional area, the effective resistance of the inner conductor is much higher than the effective resistance of the outer conductor. As the frequency of the transmitted energy is increased, the effective resistance of the conductor also increases. This is due to skin effect, that is, energy at higher frequencies tends to travel along the surface or skin of a conductor. Since this effect reduces the available cross-sectional area, it also increases the effective resistance.

The removal of the center conductor in a coaxial line eliminates it as a cause of skin-effect loss. This, of course, converts the coaxial line into a waveguide. The inner surface of the waveguide is large enough to reduce the skin-effect loss considerably.

The waveguide has a high efficiency of transmission. As a result of the three advantages listed above, the waveguide provides a very efficient transmission line for microwave energy. It eliminates the radiation loss of the two-wire line and the dielectric and inner conductor losses of the coaxial cable. Therefore, signals whose frequencies exceed 2,000 mc are usually transmitted through waveguides. However, a coaxial line is often used for frequencies up to 2,000 mc and for even higher frequencies. You will see the reason for this when the disadvantages of waveguides are discussed.

The waveguide has a high power-handling ability. In any given system of transmission, the ability to handle higher power is usually limited by the ability of the system to handle higher voltages. In a coaxial transmission line, the ability to handle high voltage is limited by the distance between the inner and outer conductors and the type of dielectric used. The larger the diameter of the line, the greater the power-handling capacity. If the diameter of the cable were too small for a given transmitted power, the voltage would arc over from the center conductor to the outer or ground conductor.

A waveguide of the same diameter, however, will handle a higher voltage than the coaxial line. You can see why this is so by referring to the illustration at upper right. Part A shows a coaxial line with a distance d from center conductor to outer conductor. As shown at B the distance d for a wave guide of the same outside diameter is over twice that of the coaxial line. The waveguide, therefore, can handle a much higher voltage and, consequently a much higher power than the coaxial line.



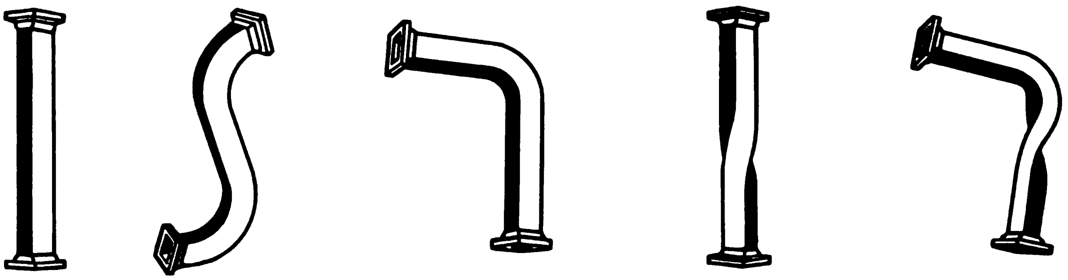
Breakdown Paths of Coaxial Line and Waveguide

Disadvantages

The coaxial line has a frequency above which its losses become too great for practical use. The waveguide, on the other hand, has a frequency below which its losses become too great for practical use. In the waveguide, the frequency at which this occurs is known as the *cutoff* frequency. At frequencies below the cutoff frequency, the attentuation in waveguides is much greater than for an equivalent coaxial line.

In general, the cutoff frequency varies inversely with the dimensions of the waveguide. This relationship is such that waveguides become practical only at microwave frequencies. For example, in order to transmit energy at a frequency of 300 mc, a 2- or 3-foot waveguide would be necessary. The physical size of such a waveguide makes its use impractical. At 3,000 mc, however, the waveguide need only be about 2 or 3 inches wide. This is a practical dimension and may be handled with relative ease. Since a waveguide passes only frequencies above the cutoff frequency, *its use is limited to the microwave frequency range* because of practical dimension considerations.

Another disadvantage of the waveguide involves its *installation problems.* The installation of waveguides presents problems that are different from those encountered with two-wire or coaxial transmission lines. Waveguide installations are often referred to as *electrical plumbing* because the exact positioning and rigidity required calls for installation techniques well known to plumbers and less known to electronic or microwave technicians.



Waveguide Sections

Special couplings, which will be discussed later, are required to fasten the lengths of the waveguide together. At the lower radio frequencies, soldering the conductors together assures adequate transmission characteristics. At microwave frequencies, precautions must be taken to prevent the leakage of any energy. Ideally, the waveguide should be one long, continuous section between transmitting and receiving points. However, the waveguide is usually made in fairly short sections which must be connected together. Some of the many waveguide sections which have been manufactured for specific applications are shown in the illustration above. Special flexible waveguide sections have also been developed to be used in those situations where the standard, or fixed, curved sections cannot be used.

You can now see why each new installation must be laid out carefully beforehand. After this has been done, the various sections of waveguide required can be ordered, and finally the whole line can be put together. Consider this type of installation in contrast to that encountered when a roll of flexible coaxial cable is used. This type of line is easily bent around corners or passed through wiring troughs. You can see then that waveguide installations are more complex and also more expensive.

The specialized nature of a waveguide installation increases the cost of this method of energy transmission over that required by a coaxial installation. In addition to installation costs, the price of the guide itself is high. The waveguide must meet rigid mechanical specifications in order to prevent discontinuities along the transmission path which might introduce energy losses. To minimize such discontinuities, special machining is performed on the waveguide during manufacture. This also adds to initial equipment cost. Also, in some applications, the inside of the waveguide may be silver-plated or even gold-plated to reduce the skin-effect losses. Again, this feature results in increased cost.

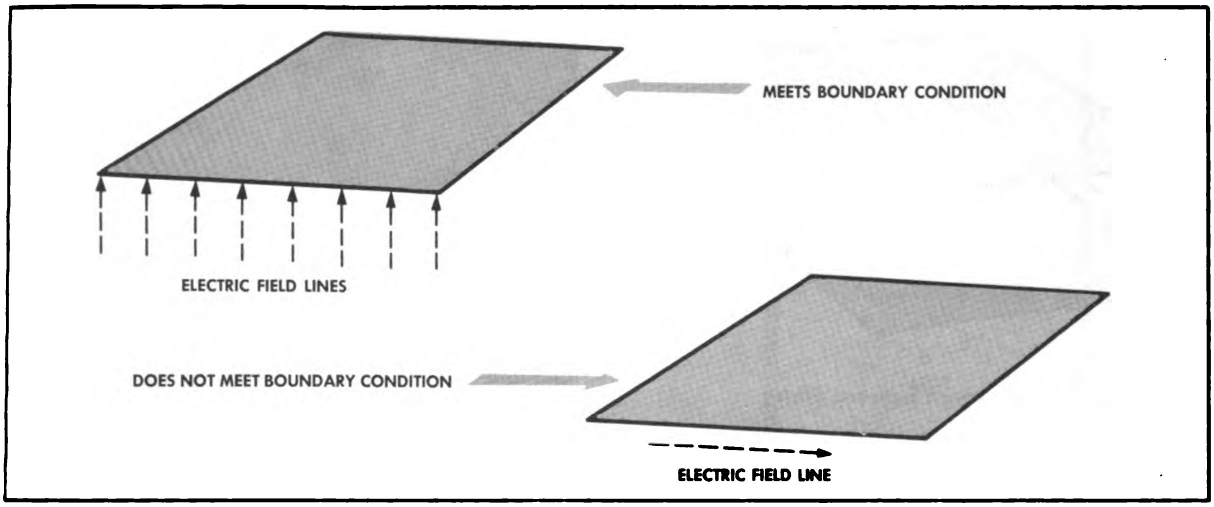
Boundary Conditions

Energy is propagated down the waveguide in much the same way as electromagnetic waves are transmitted in free space, except that these waves are kept within the boundaries of the metallic guide. However, when waves are restricted within a metallic guide, they will continue to be transmitted only if certain conditions, known as *boundary conditions*, are satisfied. If these boundary conditions are not met, the waves cannot be propagated down the waveguide.

Two boundary conditions are involved. The first is that an electric field must be perpendicular to the conductor to exist at the surface of a conductor. The converse of this boundary condition is also true. An electric field cannot exist at the surface of a conductor if it has a component that is parallel to the conductor (assuming a perfect conductor). This boundary condition is shown at the top of the next page. The second boundary condition is that a varying magnetic field must exist in closed loops parallel to the conductors and perpendicular to the electric field.

If a system meets one boundary condition, it also meets the other. It is only necessary to check for one condition to determine whether or not a wave can exist at the surface of a conductor.

Consider what happens if a plane electromagnetic wave that is radiated from an antenna

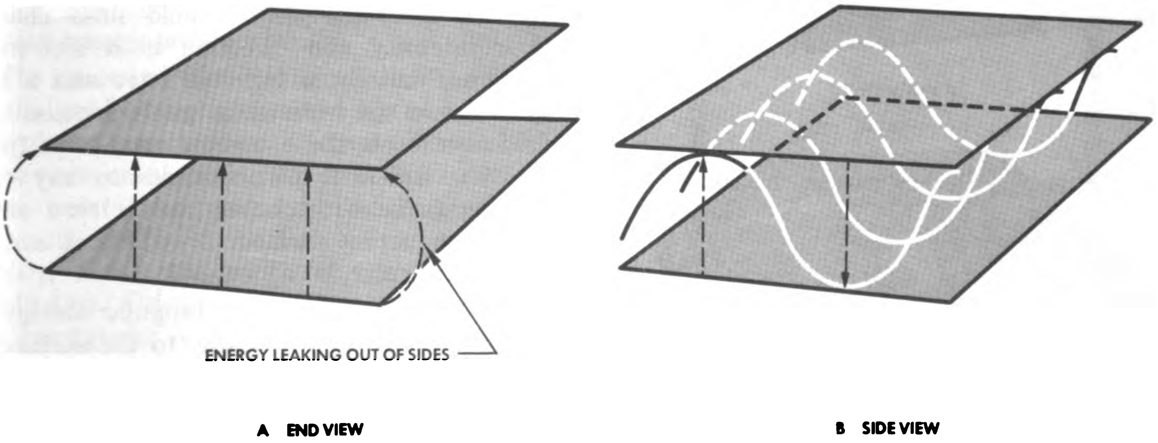


Boundary Condition for an Electric Field

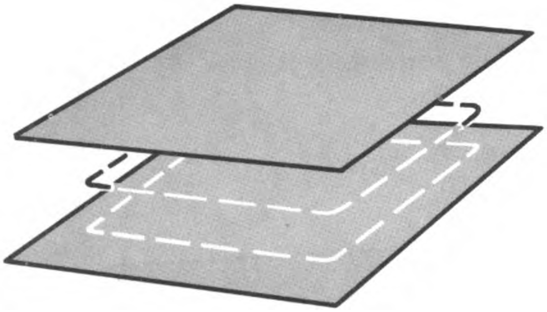
is confined between the two parallel plate conductors as shown below in the illustration, Electric Field between Plates. Obviously, both boundary conditions are met. One boundary condition requires that the electric field lines must be perpendicular to the conductors at their surfaces. This condition can be met by orienting the electric field so that the electric field (E) lines are perpendicular to the conductors as shown in the illustration. Note that this field varies sinusoidally along the distance between the plates.

A varying magnetic field accompanies the varying electric field between the two parallel plates. The magnetic field must exist in closed loops and must be parallel to the conductors. The illustration on the next page, Magnetic Loops between Plates, illustrates the magnetic field loops that exist between the conductors at a given instant of time.

Unless the width of the conducting plates is extended infinitely, some of the transmitted energy will leak or escape through the open sides. In order to avoid this loss of energy, side walls are added to the parallel plates. The result is a rectangular waveguide. The plane wave shown in the illustration, Electric Field between Plates, cannot exist in this rectangular guide because this wave has an electric field parallel to the surface of the side walls.



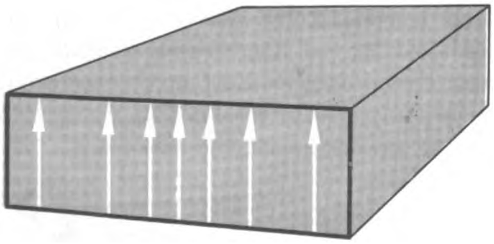
Electric Field between Plates



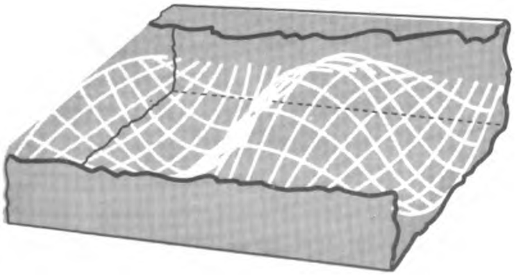
Magnetic Loops between Plates

Energy Flow in a Waveguide

There are a great many different electric field configurations that can exist in the waveguide. One of these, shown below can exist because there are no electric field lines at the side walls which are parallel to these walls. The proximity of the lines of force to each other is an indication of the strength of the electric field at that point. The field distribution shown is known as a half sine-wave across the width of the guide which is maximum at the center and zero at both sides.

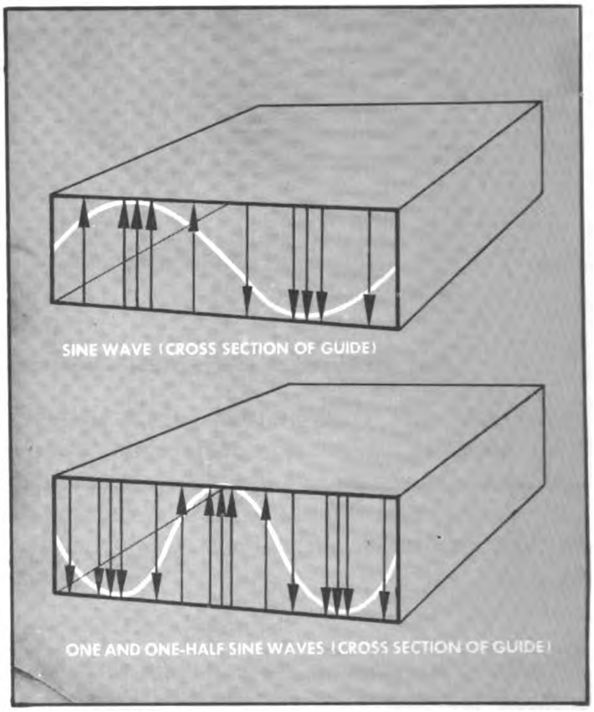


A END VIEW



B SIDE VIEW

Half Sine-wave E-field Distribution

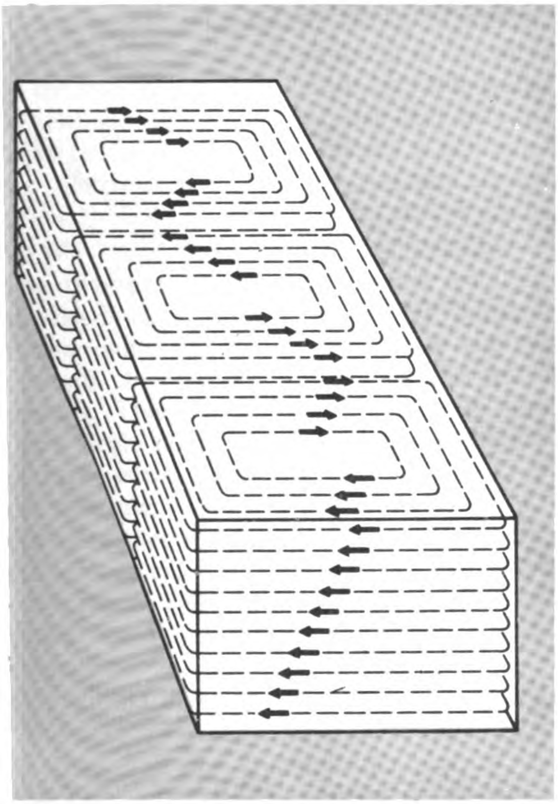

SINE WAVE (CROSS SECTION OF GUIDE)
ONE AND ONE-HALF SINE WAVES (CROSS SECTION OF GUIDE)

Other E-field Distribution in Waveguides

With the electric field moving continuously, the regions of maximum voltage move down the guide. The density of the electric field along the center of the guide increases and decreases sinusoidally. The electric field at the side walls is always zero so that the boundary condition is always met as shown in part B of the illustration.

An electric field known as a full sine-wave field can also exist in the rectangular guide because the E-field is also zero at the side walls, as shown in part A of the illustration above. In this case, the electric field lines change their intensity and direction in a sine-wave pattern. Similarly, a full sine wave and a half can exist in the rectangular guide because this field also meets the boundary conditions (part B). The E-field is perpendicular to any conducting surface it touches and is zero along those conducting surfaces to which it cannot be perpendicular.

The magnetic field in a rectangular waveguide exists in closed loops parallel to the surface of the conductor. Its strength is proportional to the rate of change of the electric field. The illustration on the next page, Magnetic Field Pattern, shows the magnetic field (H-lines) that accompanies a half sine-wave electric field

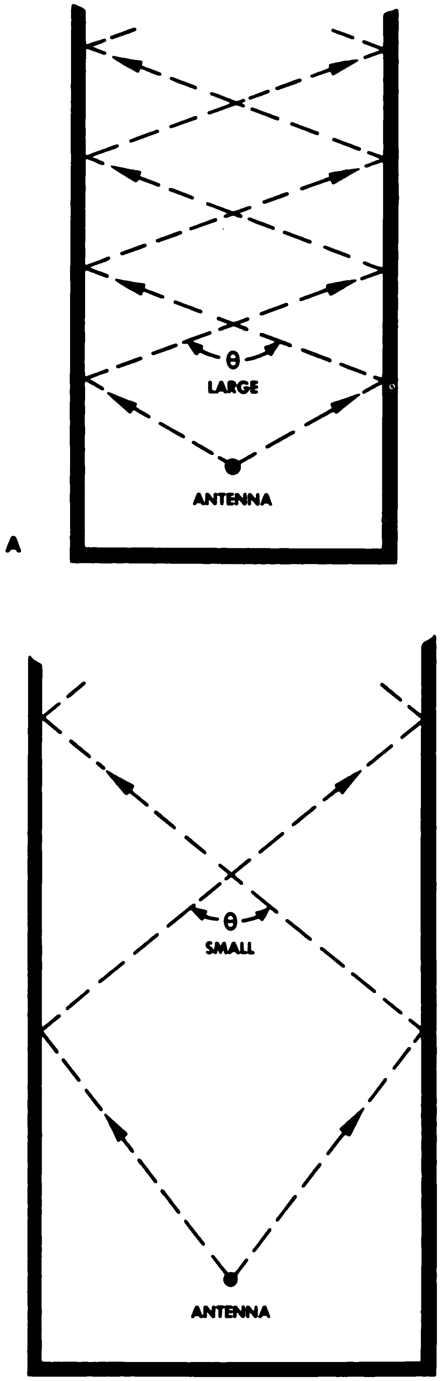


Magnetic Field Pattern

distribution. The intensity of the magnetic field varies sinusoidally down the center of the guide just as the electric field does.

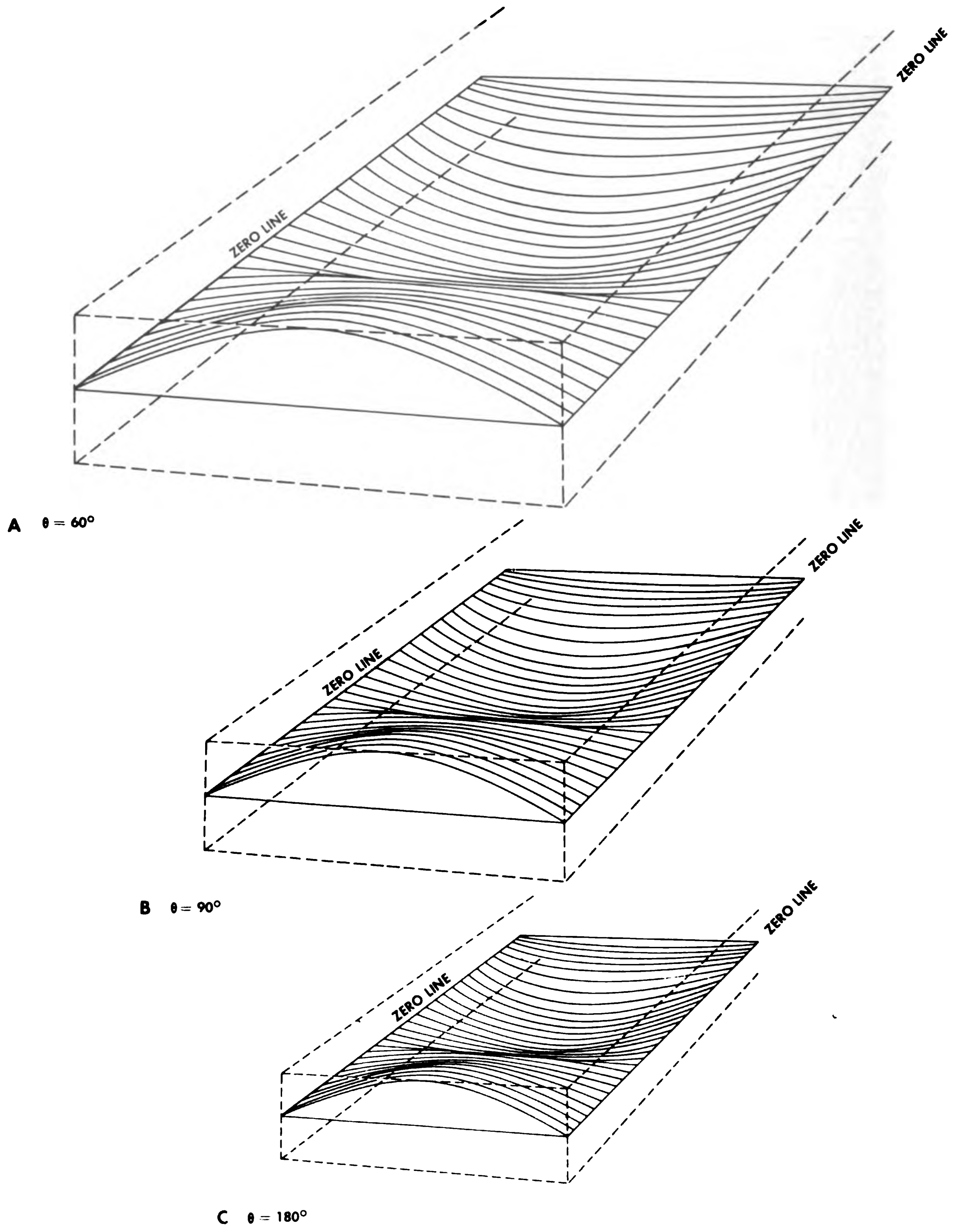
Electromagnetic waves transmitted down a waveguide are radiated into the guide by means of an antenna. This antenna radiates in all directions. However, there are only two angles of radiation at which proper addition and cancellation take place to produce a wave which fulfills the boundary conditions required to sustain energy propagation.

The sinusoidal configurations across the width of the guide can be obtained by crisscrossing two plane waves at a certain angle. When plane waves are crisscrossed, the two waves add at some points and cancel at others, and the resultant field distribution has a sine-wave pattern across the width of the guide. The illustration, Crisscrossing Plane Waves, shows electromagnetic waves propagated down the waveguide in the form of two crisscrossing plane waves. These waves bounce back and forth between the walls of the guide and cross each other at some angle *θ*. The antenna is properly oriented so that these waves strike the wall at the right incident angle to obtain the desired field configuration in the guide, fulfilling the boundary conditions.



Crisscrossing Plane Waves in the Waveguide

Note the views of the crisscrossing waves on page 388. Along the sides of the waveguide,



Resultants of Crisscrossing Waves

the two waves cancel each other. Maximum electric field intensity occurs along a line midway between and parallel to the two side walls. Furthermore, the pattern of the field perpendicular to the side walls represents a half sinewave distribution similar to the one shown previously. Consequently, the boundary conditions are met using plane-wave propagation.

The diagrams show the change in resultant field distribution when the angle θ between the two plane waves is varied. Note that as the angle θ increases, the waveguide width decreases. There is a limit, however, to how far the angle of intersection can be increased. This limit is reached when angle θ is 180°, or when one plane wave is traveling in the opposite direction with respect to the other. The width of the resultant half-sine distribution is the minimum half-sine width and is equal to one-half the wavelength of the plane wave in free space.

You should know the following facts regarding the action of these waves in waveguides:

1. The half-sine configuration can be obtained only when the width of the waveguide is equal to or greater than one-half the wavelength of the signal in free space. All signals whose wavelengths are greater than twice the width of the waveguide cannot exist in the guide.
2. If the width of the waveguide is kept constant, the angle between the two plane waves decreases as the frequency increases.
3. If the frequency is kept constant, the angle between the two plane waves decreases as the width of the waveguide increases.

Modes in Waveguides

Now that the method of energy propagation down the waveguide has been shown, some other factors can be considered. As you may have expected, some of these factors are unique to the operation of waveguides.

You have noted that a number of electric field configurations can exist or be propagated along a waveguide. Examples were given for the half, full, and one and one-half sine-wave fields. Other field configurations also meet boundary conditions and they, too, can be propagated down the waveguide. All these configurations of field patterns in waveguides are called *modes*. There are two main classes of modes. These are the *transverse electric* (TE) and the *transverse magnetic* (TM).

A transverse electric mode is one in which all components of the electric field (but not the magnetic field) lie in a plane transverse or perpendicular to the direction of propagation. In other words, with the direction of wave propagation along the Z axis, the direction of the electric lines is at right angles to the Z axis. Part A of the illustration on page 390, TE and TM Modes, shows a half sine-wave distribution of E-lines along the width of the guide. The electric field is parallel to the Y axis, and the magnetic field has both X- and Z-axis components. This mode is therefore a TE, or transverse electric mode.

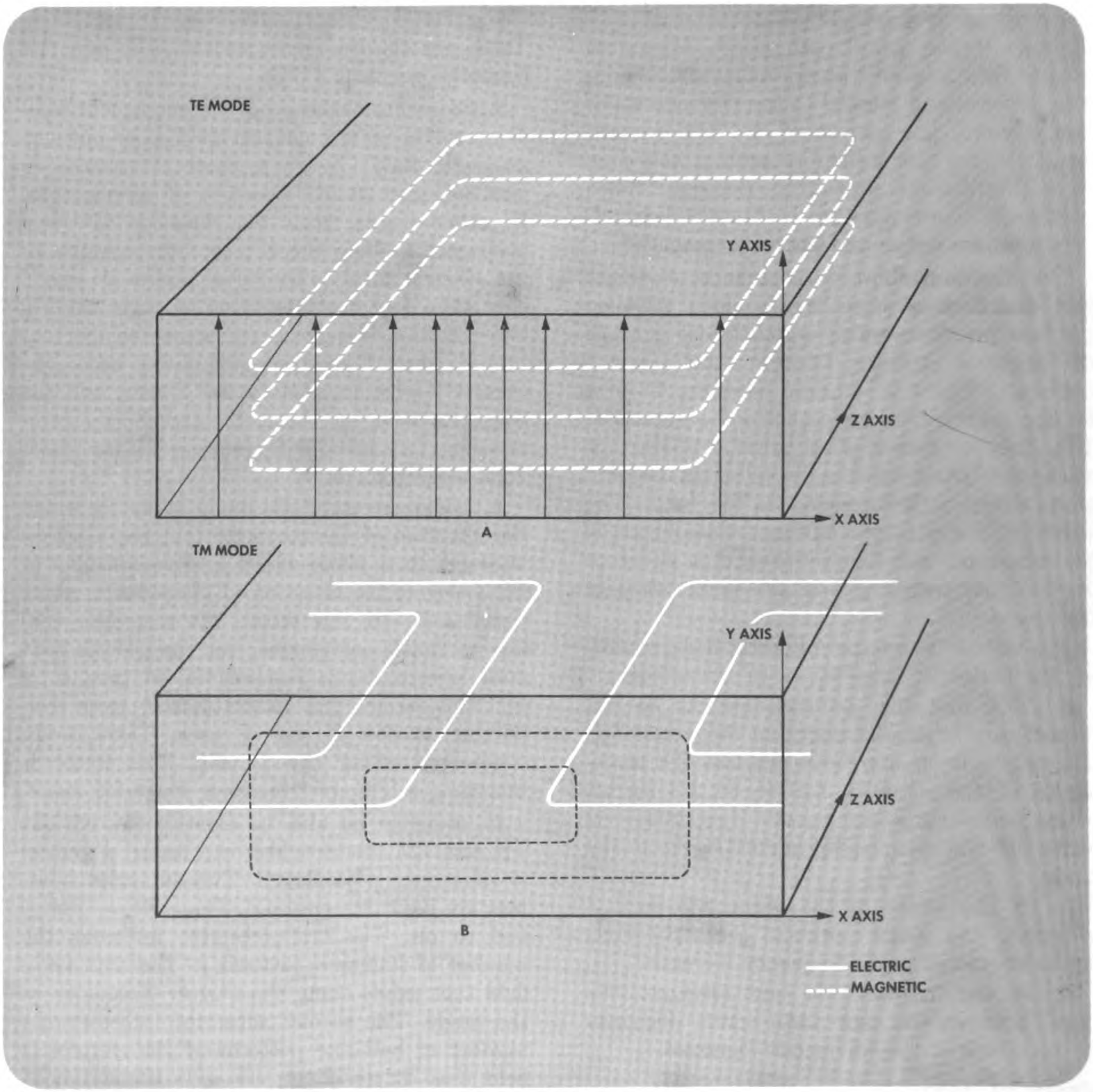
A transverse magnetic mode is one in which the direction of the magnetic (but not electric) lines lies in a plane which is perpendicular or transverse to the direction of propagation along the Z axis. In this mode, the magnetic field has no Z-axis component, but the electric field does have a Z-axis component. In part B of the illustration, the magnetic-field loops run parallel to the X and Y axes. There is no component along the Z axis. This mode is therefore, a TM, or transverse magnetic mode.

In order to be able to identify the various TE and TM modes which can occur, a system of subscripts is employed. Two subscript numbers are used, for example, $TE_{0,1}$, $TE_{1,1}$, $TM_{1,1}$, and so on. The first subscript indicates the number of half-sine patterns of the transverse field that exists along the narrow dimension of the guide. The second subscript indicates the number of half-sine patterns of the transverse field that exists along the wide dimension of the guide. When there is no change in field intensity a subscript "$_0$" is used.

For example, consider the TE mode shown in the illustration. There is no variation of the electric field along the Y axis, or narrow dimension of the guide. The first subscript number is then 0. A half-sine pattern of E-lines exists along the width, so the second subscript is 1. The field configuration is therefore called a $TE_{0,1}$ mode. The $TE_{0,1}$ mode is the one most commonly used in rectangular waveguides.

Cutoff Frequency

The electromagnetic wave can be propagated down the waveguide only when the boundary



TE and TM Modes

conditions are met. The boundary conditions can be met by crisscrossing two plane waves so that a sinusoidal distribution is obtained across the width of the guide. This can be done only if the signal frequency is higher than the *cutoff frequency*.

The cutoff frequency is different for each mode and for different waveguide shapes (such as rectangular or cylindrical). In a rectangular guide, the lowest cutoff frequency occurs for the $TE_{0,1}$ mode; the $TE_{0,1}$ mode is therefore called the dominant mode. The cutoff frequency for this mode is that frequency whose wavelength, known as the cutoff wavelength, is equal to twice the width of the waveguide. Expressed mathematically, the cutoff wavelength is,

$$\lambda_c = 2B$$

where B is the width of the guide and λ_c is the cutoff wavelength. The relationship between cutoff wavelength and cutoff frequency in a waveguide is,

$$f_c \text{ (mc)} = \frac{3 \times 10^4}{\lambda_c \text{ (cm)}}$$

Thus, if B is equal to 5cm, λ_c is 10cm. The cutoff frequency is then

$$f_c = \frac{3 \times 10^4}{10} = 3{,}000 \text{ mc}$$

Group and Phase Velocity

The speed at which a pulse of electromagnetic energy travels down the waveguide is known as the *group velocity* of the wave. This energy travels down the guide at a velocity less than that of light. The group velocity is a function of the incident angle of the two crisscrossing waves with the side walls. For large angles between the waves and the side walls (see page 388, Crisscrossing Plane Waves part A), the group velocity is reduced. In the extreme case, where the angle made by the waves with the side walls of the waveguide is 90°, the waves never go forward. They are just reflected back and forth between the walls. The group velocity is then zero, and there is no propagation of energy down the waveguide. On the other hand, the group velocity increases as the angle between the crisscrossing waves and the side walls decreases.

The group velocity is also affected by the signal frequency. Increasing the wave frequency for a fixed rectangular guide width decreases the angle. Decreasing the angle increases the group velocity of the wave. The group velocity therefore varies directly with the signal frequency. This group velocity approaches but cannot exceed the velocity of light.

In addition to the group velocity of the resultant wave in a waveguide there is another velocity known as the *apparent* or *phase velocity*. This is the velocity at which the fields seem to move down the guide. It is the product of frequency and the guide wavelength. The phase velocity in a waveguide is always greater than the speed of light. This does not violate any physical laws, since the velocity of energy propagation (group velocity) is always less than the velocity of light.

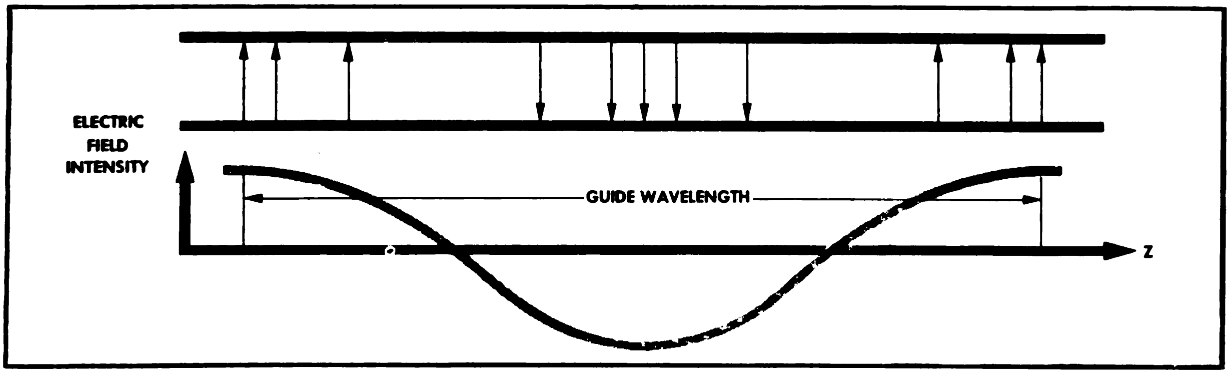
The movement of an ocean wave striking the beach at an angle is a good analogy of phase and group velocity in a waveguide. The point at which the wave strikes the beach moves very quickly from one end of the beach to the other, even though the wave front is actually moving much more slowly. The wave front may be said to move at the group velocity. The moving point on the beach may be said to move at the phase velocity.

Wavelength within the Guide

The wavelength of the resultant wave in the guide is equal to the distance between a reference point on a wave and a corresponding point in the next cycle of the wave. This is shown below in the illustration, Wavelength in Guide. The wavelength in the guide is directly proportional to the phase velocity. Since the phase velocity is always greater than the velocity of light, the wavelength within the guide is always greater than the wavelength in free space.

Characteristic Impedance

The characteristic impedance of a waveguide is the ratio between the electric and magnetic field components that are transverse to the direction of energy propagation. The characteristic impedance varies considerably for different



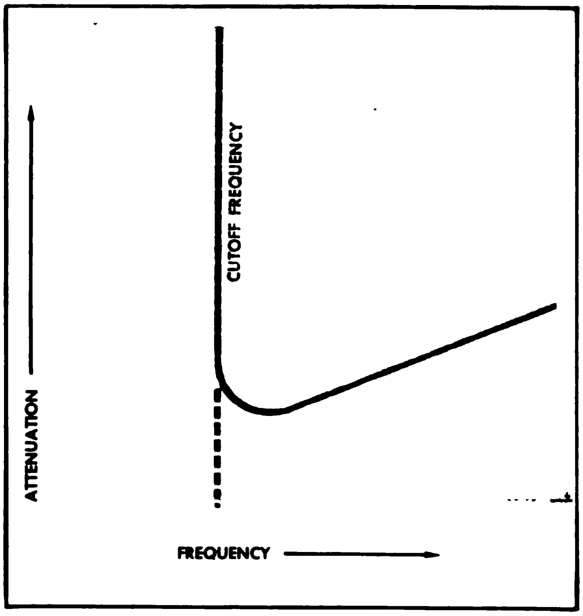
Wavelength in a Rectangular Waveguide

frequencies, modes, guide dimensions, and locations. The change of characteristic impedance with frequency should particularly be noted because it is in contrast with what happens in the other types of lines, whose characteristic impedance remains essentially constant for a very wide band of frequencies.

Attenuation

The attenuation or signal loss in a waveguide is caused mainly by currents induced in the guide walls by the traveling electromagnetic waves. These currents occur mainly at those points at which the waves strike the walls. Therefore the more often the waves strike the walls, the greater is the attenuation. Since the angle between the waves and the side walls decreases as the frequency increases, the attenuation, due to this factor, decreases with increasing frequency. Another factor affecting attenuation is the conductor resistivity, which increases with frequency.

As a result of these two factors, the attenuation versus frequency curve of a waveguide (for a particular mode of operation) generally is as shown in the illustration below. The attenuation is very high near the cutoff frequency where the angle between the crisscrossing waves and the guide walls is close to 90°. The attenuation then decreases sharply due to reduction in the angle as the frequency increases. However, a point is reached where, as the frequency is increased still more, the increase in the conductor resistivity causes an increase in the attenuation. Beyond this point the attenuation increases with frequency.



Attenuations in a Waveguide

Power-Handling Capabilities

The maximum power that can be transmitted in a waveguide depends on the maximum electric field strength that can exist without breakdown. As the charge is built up between two conductors which are separated by air, a point is reached where the air breaks down and arc-over occurs. The intensity of the electric field when the air breaks down is known as the electric field breakdown intensity or strength. If the electric field strength at any point in a waveguide exceeds this electric field breakdown strength, dielectric breakdown occurs. The microwave power is dissipated in the walls of the guide and in the arc that is formed.

The breakdown electric field strength of air varies with different pressures, moisture content, and temperature. For normal conditions, the value of the breakdown field is about 30,000 volts (peak) per centimeter. To increase the power-handling capability of radar equipment, some components are pressurized.

Common Sizes and Shapes of Waveguides

There are waveguides of many shapes and sizes that can be used to transmit microwave energy. Actually, however, only a very limited number of these shapes and sizes is used in practice. Most commercial waveguides, for example, are either rectangular or cylindrical in shape. This is true mainly because these shapes are the easiest to manufacture and the desired modes can be readily set up in them. Although there is a best waveguide size for each frequency, only a few standard sizes are employed in order to simplify manufacture, installation, and replacement.

Many modes can exist simultaneously in a waveguide, although in practice it is usually desirable to have only one mode at any one time. The antenna used to take energy out of the guide at the receiving end is often unable to pick up the power from the electromagnetic wave when modes other than the one for which the antenna was designed are used. As a result, any other modes that might exist not only fail to perform any useful function, but increase the attenuation.

Consequently, a waveguide system is designed so that only one mode is propagated down the guide. This mode, known as the dominant mode, is generally the simplest one that can exist in the guide. In rectangular guides, the dominant mode is the $TE_{0,1}$. In cylindrical guides, the dominant mode is the $TE_{1,1}$.

The tables shown below list some of the standard waveguide sizes, typical cutoff frequencies, usable frequency bands, and attenuations at specific frequencies for both rectangular and cylindrical waveguides.

Coupling Methods

In a two-wire, or coaxial-line, transmission system, the generator is coupled to the line by the connection of its output terminals directly to the line conductors. With waveguides, this method cannot be used because there are no comparable conductors that can be connected to the generator terminals. Instead, the generator energy must be transformed into an electromagnetic wave by means of an element that acts as an antenna. This is then the method of coupling the generator energy into the guide. Similarly at the output end of the guide, the electromagnetic energy must be extracted from

RECTANGULAR WAVEGUIDES ($TE_{0,1}$ MODE)

Dimensions in inches	Cutoff Freq. in mc	Usable Freq. Band in mc	Attenuation in brass waveguide-db/ft
1.5 x 3	2,074	2,540-3,950	0.012 at 3,000 mc
1 x 2	3,154	3,950-5,840	0.021 at 5,000 mc
0.75 x 1.5	4,304	5,840-8,100	0.036 at 6,000 mc
0.625 x 1.25	5,263	7,050-10,000	0.050 at 8,300 mc
0.5 x 1	6,562	8,100-12,500	0.076 at 9,400 mc

CYLINDRICAL WAVEGUIDES ($TM_{0,1}$ MODE)

Diameter in inches	Cutoff Freq. in mc	Usable Freq. Band in mc	Attenuation in (brass waveguide db/ft)
4.0	2,260	2,750-3,400	0.002 at 3,000 mc
3.0	3,010	3,800-4,300	0.003 at 4,000 mc
2.75	3,400	4,000-4,760	0.0037 at 4,500 mc
2.5	3,620	4,700-5,250	0.004 at 5,000 mc
1.193	7,600	9,000-11,000	0.014 at 10,000 mc
0.933	9,650	11,000-14,000	0.208 at 12,000 mc

the guide. This is known as coupling energy out of the guide.

There are three basic methods of coupling energy into and out of a wave guide. They are the *probe*, the *loop*, and the *window* (also called iris or slot) methods.

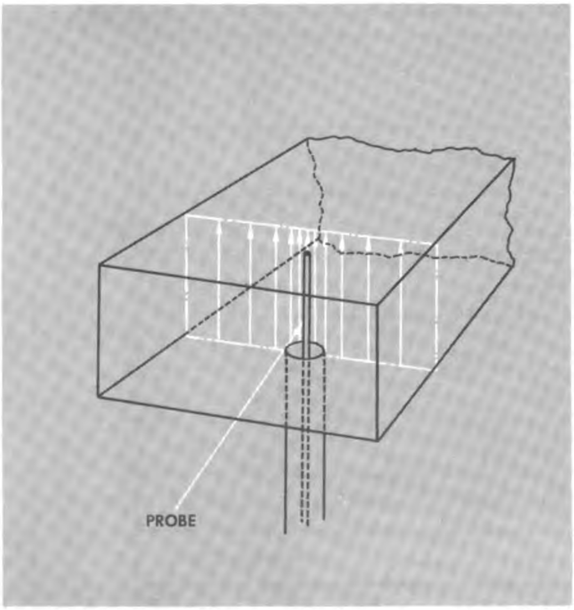
PROBE COUPLING. A coupling probe is a small metallic conductor which is inserted in the waveguide parallel to where the lines of the electric field are to occur. When it is used for coupling energy into the guide, the probe sets up electric field lines that the guide can support. The probe shown in the illustration below is inserted in a guide to be operated in a $TE_{0,1}$ mode. For this mode, the electric field is maximum at the center of the width of the guide, and hence the probe is placed at this position. When the probe is used to couple energy out of the guide, it is placed in a similar position. A current caused by the presence of the changing field is then induced in the probe. The energy thus induced in the probe is transferred to the load.

Maximum coupling is obtained by inserting the probe at a point of maximum electric field intensity. If less coupling is desired, the probe may be placed at a point in the guide where the electric field is less intense. More or less coupling may also be obtained by varying the length of the probe. Usually the probe is made about a quarter of a wavelength long so that its action is somewhat similar to that of a quarter-wave antenna.

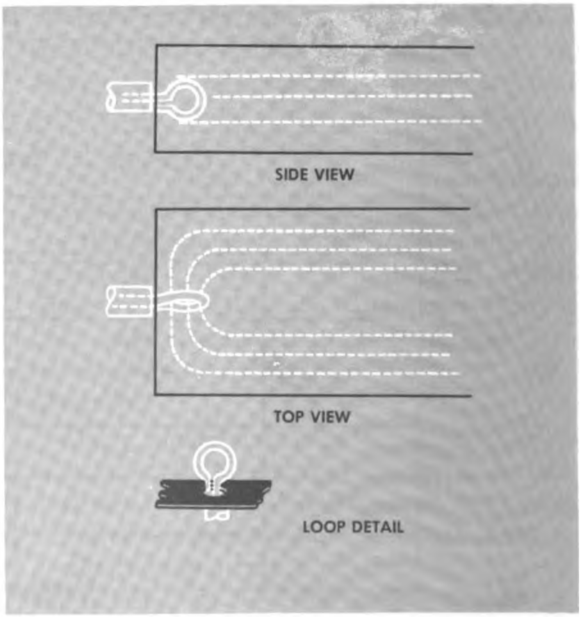
LOOP COUPLING. The illustration, Loop Coupling, shows the manner in which a loop is inserted in the waveguide to couple energy into or out of the guide. As you see in this illustration, a small loop of wire is used to link the magnetic field lines. The voltage induced in the loop is proportional to the number of magnetic flux lines that cut it.

To extract maximum energy from the guide, or to insert energy with maximum efficiency, the loop is placed at a point of maximum magnetic field intensity. It is then oriented in such a manner that the magnetic field lines are perpendicular to the loop's plane. If less coupling is desired, the loop may be rotated so that fewer magnetic lines link it. When the loop is rotated 90° from the maximum coupling position, no magnetic lines link the loop. This position of the loop results in minimum coupling. Moving the loop to a point where there is less magnetic field intensity is another method of reducing the coupling.

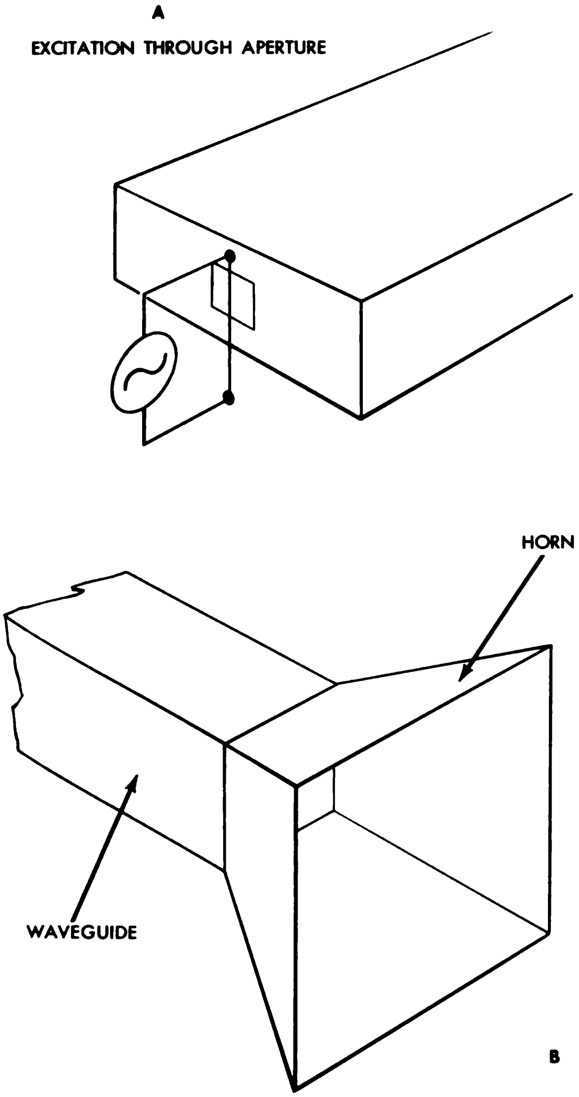
WINDOW, IRIS, OR SLOT COUPLING. A third method of coupling energy into or out of a waveguide is through a slot or opening in the guide wall. The slot allows electromagnetic energy to



Probe Coupling



Loop Coupling



Coupling through Opening in Waveguide

leak out of the waveguide. The coupling may be increased or decreased by varying the size and location of the opening. This method is sometimes used when very loose coupling is desired. In this method energy enters the guide through a small aperture, as you can see at A in the illustration above. Any device which will generate an E-field may be placed near the aperture and the E-field will expand into the waveguide. If the frequency is correct and the size of the aperture properly proportioned, energy is transferred to the waveguide with a minimum of reflections.

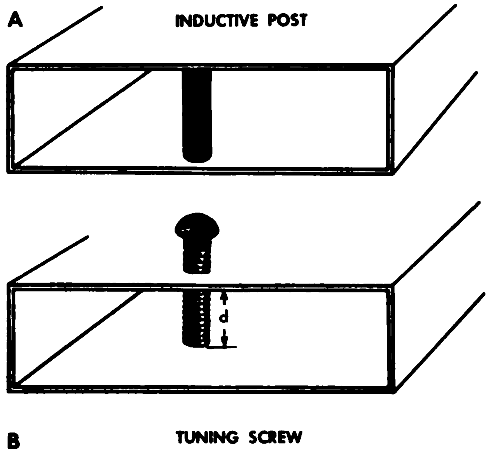
In some cases the opening is flared into a horn and used to radiate energy into space like an antenna as shown at B in the illustration to the left. This makes the guide similar to a V-type antenna. The funnel, in effect, eliminates reflection by matching the impedance of free space to the impedance of the waveguide. When the mouth of the funnel is exposed to electromagnetic fields, they enter the funnel where they are gradually shaped to fit the waveguide. The funnel is directional in characteristic. It sends or receives the greatest amount of energy from in front of the opening.

Impedance-Matching Techniques

When you couple energy into or out of a waveguide, you should effect a proper impedance match between the generator, the coupling elements, the waveguide, and the load. In this respect, the waveguide, acts exactly like a two-wire line and follows the same principles outlined previously. In other words, unless you terminate the waveguide in its characteristic impedance, standing waves are created.

Generally you used appropriate matching elements, such as impedance-matching transformer and reactive elements, to effect maximum transfer of energy. While the same principles are used in matching that were described earlier, the physical appearance of the matching elements is somewhat different in waveguides.

Post or tuning screw. You can obtain a shunt reactance across a waveguide by inserting a metallic element which may be cylindrical in shape or in the form of a strip When this element extends across the full height of the guide as in part A of the illustration below, it is called



Waveguide Reactances

an inductive post. When the element is made in the form of a screw so that the depth of insertion into the guide can be varied, as shown in part B, it is called a tuning screw.

The reactance obtained by a tuning screw varies with the depth of insertion d. When d is small, the screw acts as a shunting capacitance. Increasing d increases the shunting capacitance until the point where d becomes equal to a quarter of a wavelength. At this point, resonance occurs and the reactance is zero. For greater values of d, the screw becomes inductive. The height of the guide is usually greater than one-quarter wavelength so that, when the element extends across the full height, it is inductive in nature. In this case it is referred to as an inductive post as shown at A.

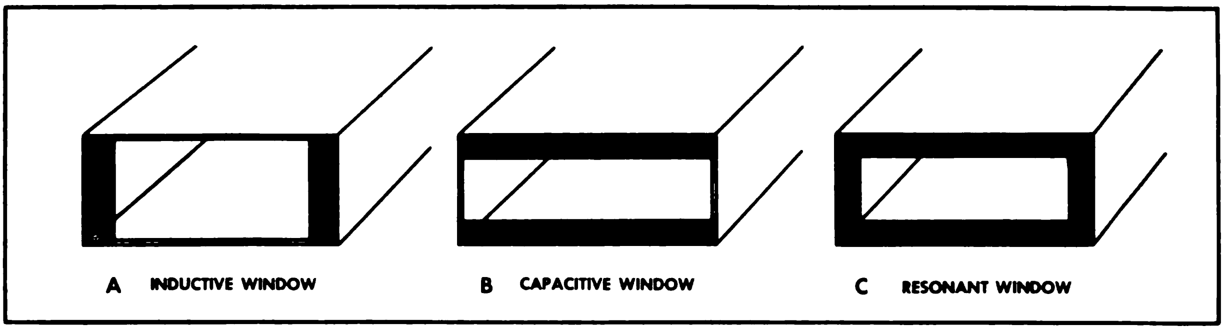
WINDOW REACTANCES. Another method of obtaining a shunt reactance is to place metal strips or partitions in the guide at right angles to the direction of energy propagation. When the strips are placed as in part A of the illustration, Window Reactances, an inductive reactance is obtained and the element is called an inductive window. The value of inductance decreases as the width of the strips increases. When the metallic strips are placed as shown in part B, the element is known as a capacitive window. The value of capacitance decreases as the width of the strips increases. When the metallic elements are placed along all walls, as shown in part C, there is both a shunt inductance and a shunt capacitance across the guide. This type of window, called a resonant window, or diaphragm is the equivalent of a parallel tuned circuit. At the resonant frequency, the inductive reactance is equal to the capacitive reactance and the window represents a pure resistance.

STUB REACTANCE. Another method of obtaining a reactance is to use an open or short-circuited stub in much the same manner as stubs are used in lower frequency lines. Shorted stubs are almost always used to prevent undesired radiation of energy. When the stub is placed as shown in part A of the illustration on the next page, Waveguide Stubs, it acts as an impedance in series with the line and is called a series stub. When the stub is inserted as shown in part B, it acts as a shunt impedance across the guide. This type of stub is called a shunt stub.

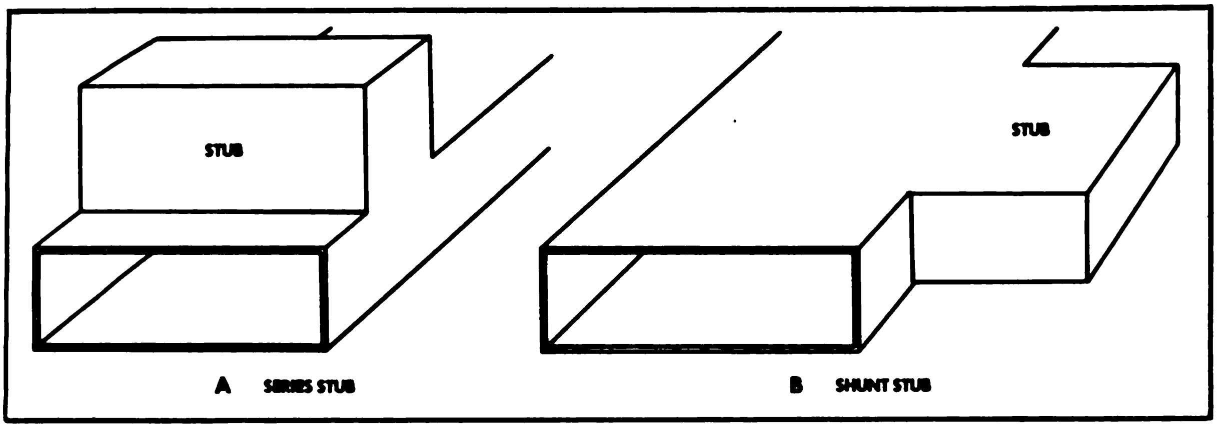
IMPEDANCE TRANSFORMERS. As mentioned previously, an impedance transformer is used to change one value of resistance to another. This transformer is placed at a point along the waveguide where the impedance is pure resistance, but not equal to Z_0. The impedance transformer then transforms this value of resistance to Z_0 and thereby matches the waveguide to the load.

You have seen that it is possible to match any load Z_R by a section of transmission line one-quarter wavelength long with a characteristic impedance equal to $\sqrt{Z_i\ Z_R}$. A *quarter-wave* transformer can be made by using a section of waveguide whose dimensions are either greater or smaller than those of the main waveguide, as shown in part A of the illustration at the bottom of the next page.

Another type of transformer that is widely used is the *tapered-line* transformer shown in part B of the illustration. Using this transformer, the dimensions of the guide are varied very gradually, thereby changing its characteristic impedance from that of the main line to the value of the load. When this change in impedance is made very gradually and the tapered section is longer than two wavelengths, the impedance



Window Reactances



Waveguide Stubs

transformation is effective over a very wide band of frequencies. In contrast, the quarter-wave transformer is effective only for those frequencies at which the line is of proper length.

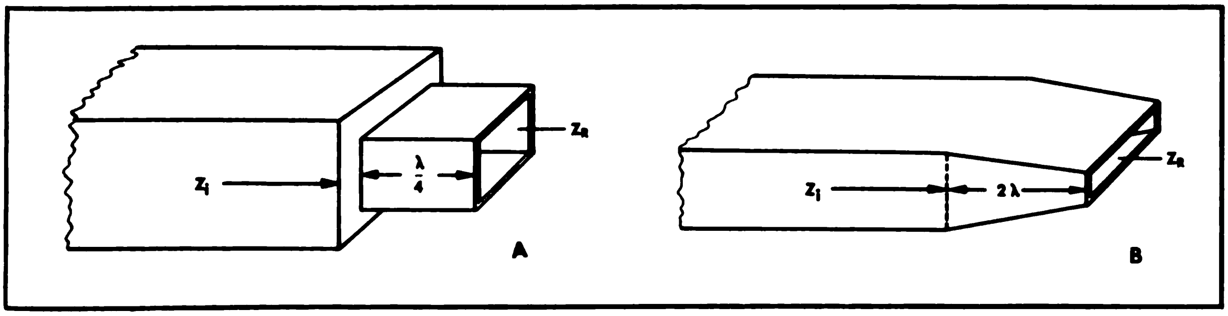
Bends and Twists

In some applications of microwave equipment, it is necessary to change the orientation of the waveguide, as shown in part A of the illustration at the top of page 398. The orientation of a waveguide is changed to effect a change in the polarization of the fields for matching purposes or to accommodate a special type of load. The basic principle involved in a twisting waveguide is to make the change as gradual as possible. As in the case of the tapered line, a gradual change serves to minimize reflections.

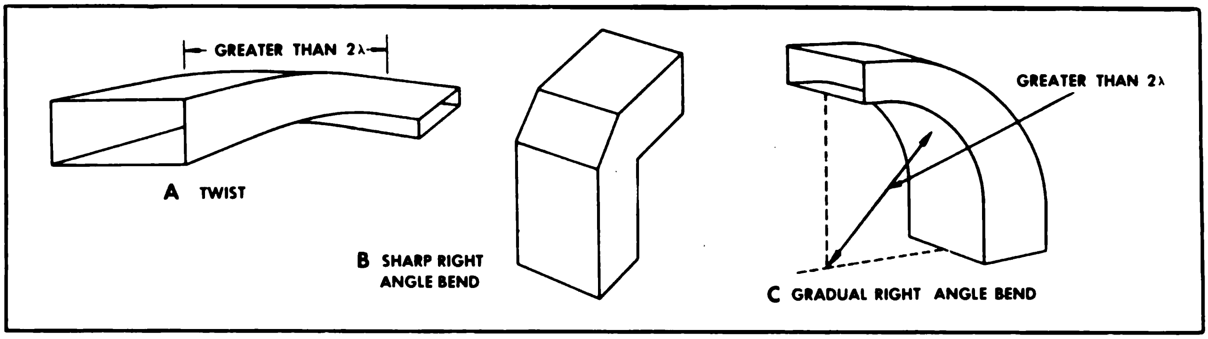
Waveguide bends or elbows are used for making right-angle turns. There are several types of bends in use. One of these is a sharp right-angle bend, also called a corner, shown in part B. Another is a more gradual bend as shown in part C. In the sharp bend, the top wall of the guide is beveled to minimize reflections. In the more gradual bend, reflections are minimized by making the radius of the bend at least two wavelengths long. This gradual change produces a very small mismatch and therefore reduces reflection losses.

Connection Methods

The simplest method of connecting two sections of waveguide together is the *flange-to-flange*, or *contact*, coupling shown in the illustration on page 398. In order to assure good electrical contact, the surfaces to be joined should be carefully cleaned to remove corrosion. By moving the flange back a short distance from the end of the guide, as shown in the illustration, a greater pressure may be applied between the two guides to assure a good contact over the entire surface. The efficiency of such a joint is affected by the tightness of the connection and the accuracy of the machining of the end surfaces.



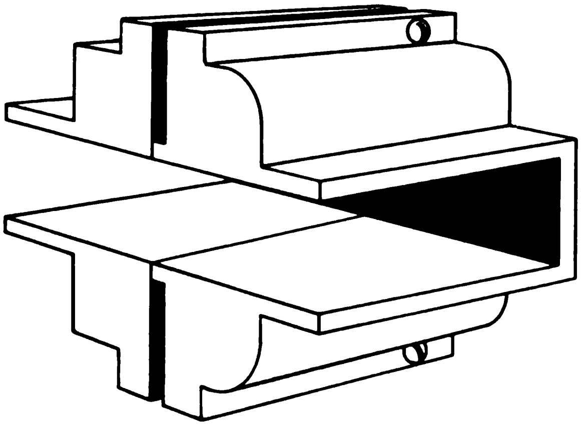
Waveguide Transformers



Waveguide Bends and Twists

A *choke joint* is a joint which provides a good electrical microwave connection between sections of waveguide, but which does not require contacting metal surfaces. Cross-sectional views of two choke joints are shown at A and B in the illustration on the next page. When the joint shown at A is used, the two sections of waveguide are bolted together. A pressure gasket is used to seal the joint and keep moisture out of the guides. When the joint shown at B is used, the sections of waveguide are actually separated by a small distance. This latter type of choke joint is sometimes called a wobble joint, since one of the sections of waveguide can move slightly, or wobble, with respect to the other.

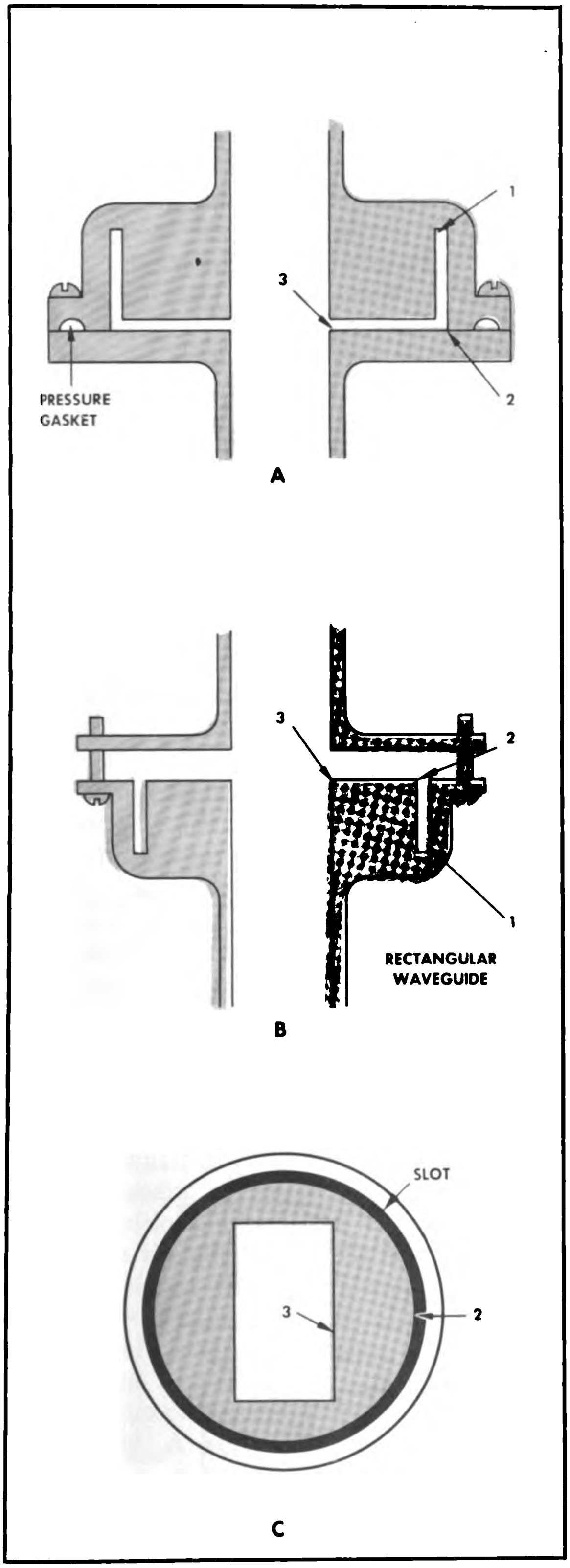
In the front view of the choke flange shown at C, you can see that it has a circular slot cut into it. The depth of this slot (between points 1 and 2) is one quarter wavelength. The flange is designed so that the middle of the broad face of the waveguide is a quarter wavelength from the edge of the slot (between points 2 and 3). Because the slot is short-circuited at point 1, there is a high impedance at point 2. This high impedance at point 2 reflects a low impedance at point 3, a quarter-wavelength away. In effect, therefore, there is an electrical connection between the two guides at point 3 for those frequencies at which the distance between 1 and 3 represents one-half wavelength.



Flange to Flange Coupling

Rotating joints are frequently used in waveguide systems which terminate in antennas that must be rotated. To permit one section of cylindrical waveguide to rotate with respect to the other, a choke joint of the wobble type can be used. As was shown previously, no actual mechanical contact is needed in this type of joint. Instead, the electrical connection is made because of the low impedance that exists between the two sections of waveguide.

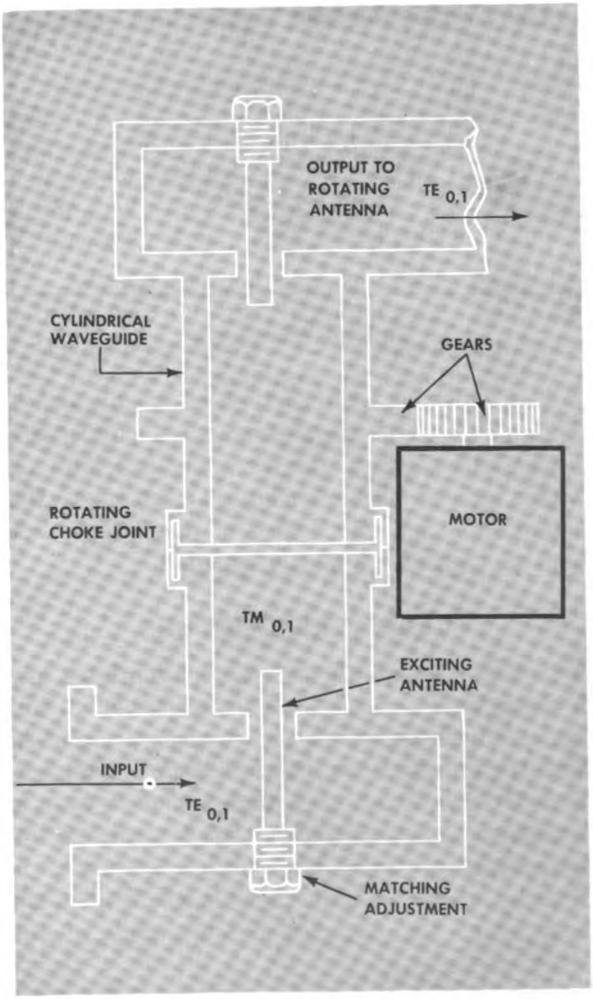
A complete system utilizing rectangular and circular waveguides is shown in the illustration at the right on page 399. In this case, a rectangular guide is used to transfer the energy from a microwave oscillator to a rotating antenna. The energy must be transferred from the rectangular guide to a cylindrical guide. Rotation of the waveguide occurs at the choke joint located at the midpoint of the cylindrical section. The energy is then transferred back to a rectangular guide, which feeds the antenna. Note that the entire top assembly above the choke joint rotates as one unit.



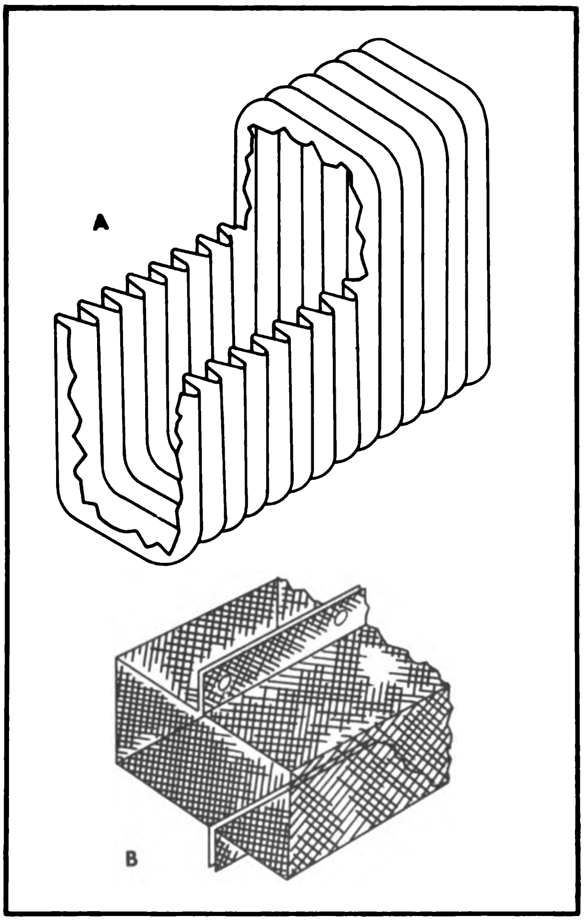
Choke Joints

Flexible Waveguides

In some waveguide applications, there arises a need for a rectangular waveguide that is flexible. This may be used for slight bends between two fixed guides or it may be used for twist and bend applications. On page 400, part A of the illustration, Flexible Waveguides, illustrates one type of flexible waveguide. The guide is made by winding a continuous metal strip in a spiral around a rectangular form. The edges of the strip are interlocked and tightly crimped together to make a reliable electrical contact. Sometimes the guide is encased in a synthetic rubber molding. With this protection, the flexible waveguide may be flexed a great deal mechanically with little change in its electrical properties.



Rotating Joint



Flexible Waveguides

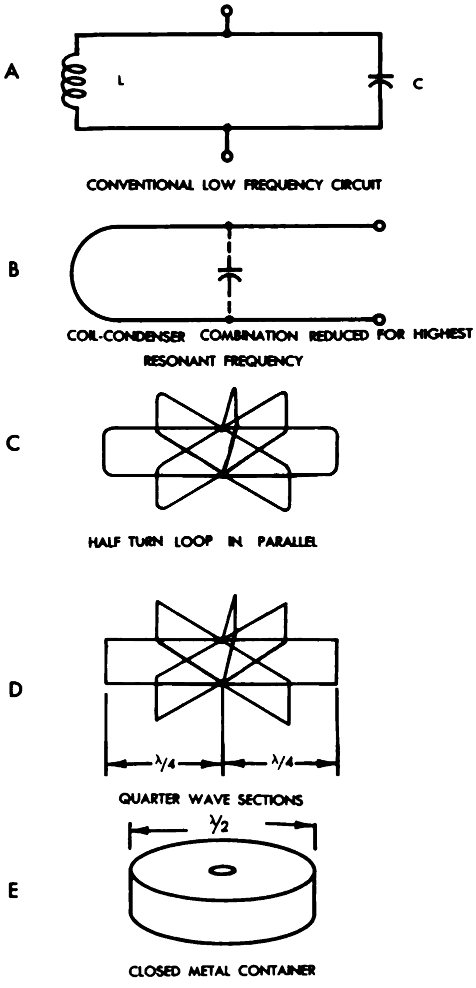
Part B shows another type of flexible waveguide. This type is constructed of copperwire screen. Here, again, the screening is reinforced with a molded rubber casing. In both types, the attenuation and the degree of mismatch when bent or twisted are relatively small.

CAVITY RESONATORS

In ordinary radio work, the conventional low frequency resonant circuit consists of a coil and a condenser which are connected either in series or in parallel as shown at A in the illustration. To increase the resonant frequency, it is necessary to decrease either the capacity or the inductance or both. However, a frequency is reached where the inductance is a half turn coil and where the capacity consists only of the stray capacity in the coil as shown at B. At extremely high frequencies this resonant circuit consists of a coil about an inch long and a quarter inch across. In this circuit the current-handling capacity and breakdown voltage for the spacing would be low.

The current-carrying ability of a resonant circuit may be increased by adding half turn loops in parallel. This does not change the resonant frequency appreciably because it adds capacity in parallel — which lowers the frequency, and adds inductance in parallel — which increases the frequency. As the effects of each cancel, the frequency remains about the same.

At C several half turn loops are added in parallel. At D several quarter wave Lecher Lines, which you recall are resonant when they are near a quarter wavelength, are shown in parallel. When more and more loops are added


A
L
C
CONVENTIONAL LOW FREQUENCY CIRCUIT
B
COIL-CONDENSER COMBINATION REDUCED FOR HIGHEST RESONANT FREQUENCY
C
HALF TURN LOOP IN PARALLEL
D
λ/4
λ/4
QUARTER WAVE SECTIONS
λ/2
E
CLOSED METAL CONTAINER

Development of Resonant Cavity

in parallel, the assembly eventually becomes a closed resonant box as shown at E, which is a quarter wave in radius or, in other words, a half wave in diameter. This box is called a *resonant cavity.*

A resonant cavity displays the same resonant characteristics as a tuned circuit composed of a coil and condenser. In it there is a large number of current paths. This means that the resistance of the box to current flow is very low and that the Q of the resonant circuit is very high. While it is difficult to attain a Q of several hundred in a coil of wire, it is fairly easy to construct a resonant cavity with a Q of many thousand. Although a cavity is as efficient at low frequencies as at high frequencies, the large size required at low frequencies prohibits its use at those frequencies. For example, at one megacycle, a resonant cavity would be a cylinder of about 500 feet in diameter. When the frequency is in the vicinity of 10,000 megacycles, the diameter of the cavity is only 0.6 inch. This makes the cavity smaller than a conventional tuned circuit. Therefore, equipment which operates at a frequency of 3000 mc or above usually employs resonant cavities as resonant circuits.

If there were no losses in the cavity, the waves would bounce back and forth between the end walls indefinitely. However, just as in any resonant circuit, some energy is lost during each cycle. Usually, a source of microwave energy feeds the cavity continuously so as to overcome the losses.

Q of a Resonant Cavity

An important characteristic of any tuned circuit, whether it is an ordinary LC circuit or a resonant cavity, is its Q. The circuit Q is the ratio between the amount of energy stored in the circuit and the amount of energy lost. A circuit with low losses has a high Q. Thus, the Q is an indication of the efficiency of the tuned circuit. The circuit Q also gives an indication as to the sharpness of tuning and the amount of impedance change off the resonant frequency. A high Q means that the circuit impedance changes a great deal off resonance, with a consequent narrow bandwidth. A low Q means that the circuit impedance changes very little off resonance, with a resultant broad bandwidth.

A cavity resonator has two values of Q. One is known as the unloaded Q, which is obtained when no energy is extracted from the cavity. The unloaded Q depends on the mode of operation and the size, shape, and construction of the cavity. Typical values range between 15,000 and 20,000 — values that are impossible to approach using conventional LC circuits.

The second value is called the loaded Q. This is the value taken with the cavity delivering energy to a load. In this case, the power lost per cycle is naturally greater and therefore the Q is greatly reduced. The actual value of Q depends on the load on the cavity. In any case, the cavity Q is usually much higher than could be obtained with conventional LC circuits.

Method of Tuning

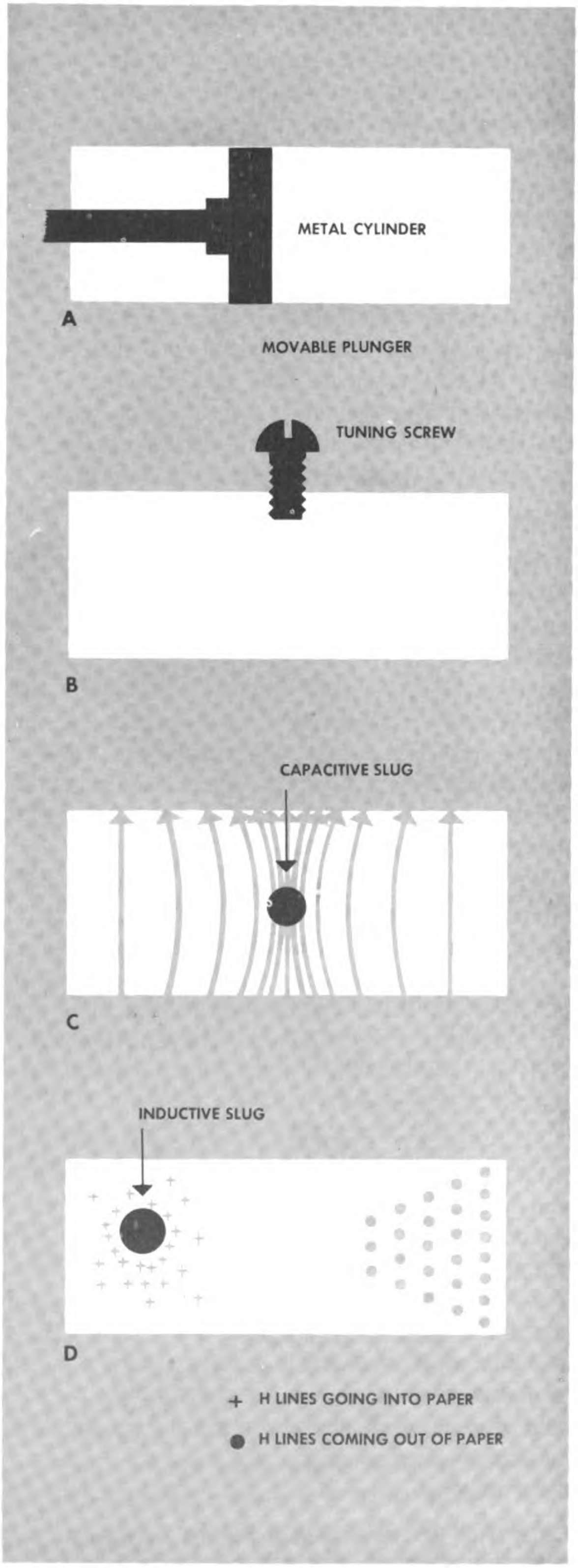
You might want to tune the resonant cavity. There are several methods of accomplishing the tuning. One method of tuning a cavity to a new resonant frequency is to vary the effective length with a plunger. This method is shown in part A of the illustration on the next page. In the simple case illustrated, the position of the end wall of the cavity is changed and oscillations at different frequencies occur as the plunger is moved back and forth.

Another method used to tune a cavity is to introduce a tuning screw into it. Essentially, the screw acts as a variable capacitance inserted at a point of maximum electrical field strength, as shown at B. As the capacitance is increased by decreasing the spacing between the screw and the wall of the cavity, the resonant frequency is decreased.

Still another method of tuning a resonant cavity is to introduce a metal slug into it. When you place the slug at a point of maximum electric field intensity, as shown at C, you shorten and intensify the electric field lines. You thus increase the effective capacitance and lower the frequency. If you place the slug at a point of maximum magnetic field strength, as shown in part D, you lengthen the magnetic field lines. You thereby decrease the effective inductance and increase the resonant frequency of the cavity.

Coupling Energy in and out of Cavities

In all of the applications of resonant cavities there must be some means of feeding energy into or extracting energy from them. Because resonant cavities are actually resonant sections of waveguide, the same methods of coupling can be used as those used with the waveguide. These include the probe, loop, and window coupling methods discussed on pages 394-395.



Methods of Tuning a Resonant Cavity

Practical Applications

The most obvious application of the resonant cavity is as a high-Q tank circuit for use at a microwave frequency. With the cavity employed as a tank circuit for a microwave oscillator, a stable and reliable microwave signal source is obtained.

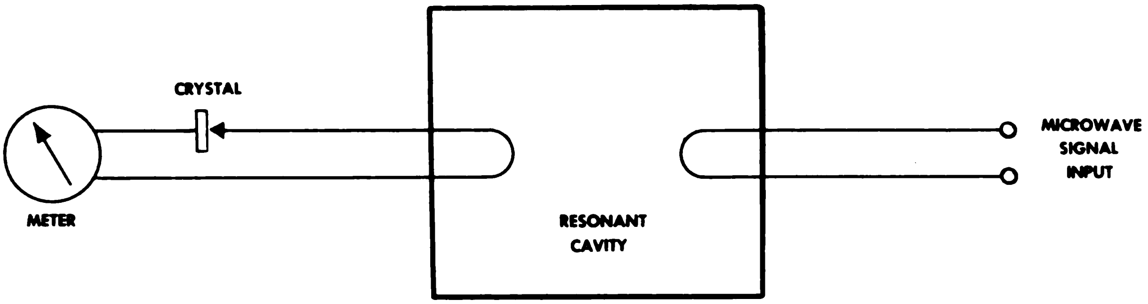
Another application of a resonant cavity is as a means for determining the over-all system performance of a radar set. In this application the cavity is called an echo box. The echo box is a high-Q resonant cavity which is used to store energy received from the main pulse of the radar transmitter. It then delivers the stored energy to the radar receiver after the main pulse has ended.

Another application of the cavity resonator is as a microwave wavemeter. A typical cavity wavemeter circuit is shown in the illustration at the top of the next page. The operation of this type of wavemeter depends on the fact that the maximum power transmitted through a cavity is obtained at resonance. Thus, the indicating meter will read a maximum when the cavity tuning adjustment is such that the resonant frequency of the wavemeter cavity is the same as that of the signal being measured. If the cavity frequency is accurately calibrated, a highly selective and accurate means of measuring microwave frequencies is obtained.

SIMULTANEOUS TRANSMISSION AND RECEPTION WITH COMMON ANTENNA

In most applications, you must transmit and receive RF signals. Two separate antennas and transmission lines are sometimes used, one for each function. A considerable simplification in equipment and installation can often be effected if the same transmission line and antenna can be made to serve both purposes.

Simultaneous transmission and reception over one antenna system can be accomplished by suitable modifications of the transmission line. Two basic methods are involved. One is used



Cavity Wavemeter

when the transmitter and receiver operate on the same frequency, as in radar. Both systems must act to keep the transmitter signal out of the receiver.

This is very important since the high transmitter power could either damage the sensitive receiver or desensitize it to a point where it could not receive the desired signal.

Parallel Duplexing Systems

In radar applications, where both receiver and transmitter are on the same frequency, single antenna operation is effected by a duplexer system which contains a TR and ATR circuit. The TR circuit prevents the transmitter power from reaching the receiver. The ATR Circuit serves to disconnect the transmitter from the antenna during reception periods.

Consider the duplexer system shown in the illustration on the next page. In this case the transmission line is drawn as a two-wire line. However, it may be either a coaxial line or a waveguide. In the case of airborne radar it is a waveguide. All operate on the same principle, and the illustration can represent any transmission line used.

In the system shown, the receiver transmission line is placed in parallel with the transmission line to the transmitter. A spark gap is in the receiver line, placed a quarter wavelength from the junction. An additional quarter-wave line is placed at a distance of one quarter wavelength from the receiver line junction toward the transmitter. It is also placed in parallel with the transmitter and is terminated in a spark gap.

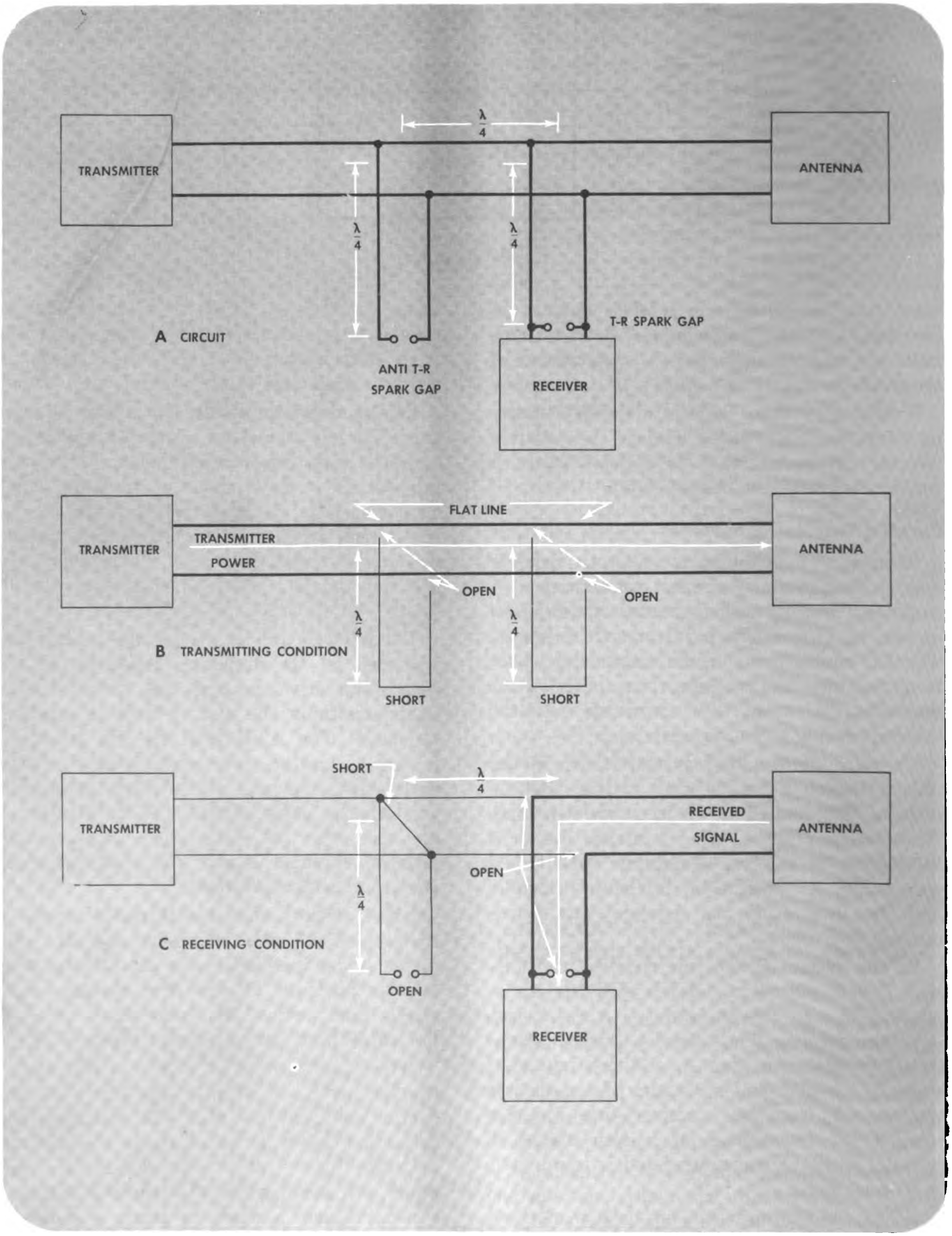
When the transmitter is on, both spark gaps fire, as shown at B. This causes both the TR and ATR circuit to act as shorted quarter-wave transmission lines. They reflect open circuits, maximum impedance, at their junction to the main transmission line. Thus the transmitter power is conducted to the antenna without loss.

During reception, neither gap is fired. Because the ATR gap is located a quarter wavelength from the main transmission line, a short circuit is produced at its T junction, as shown at C. This results in an open circuit across the receiver line, which is connected another quarter wavelength away. Therefore, the receiver pulse is conducted into the receiver section without loss.

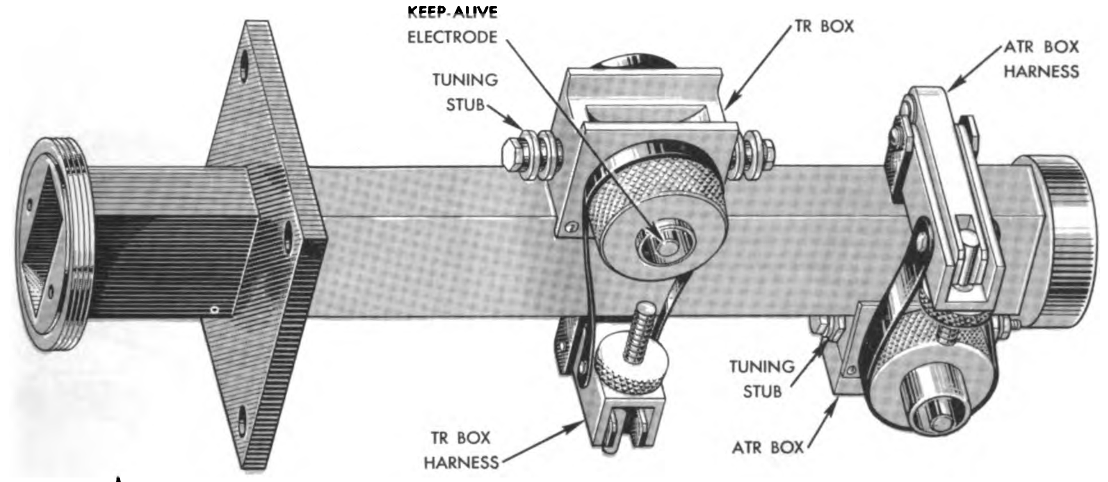
The APQ-13 utilizes the parallel method just discussed in its duplexer. Referring to the illustration shown on page 405, notice that we have essentially the same circuit as was just discussed. The ATR and the TR tubes are placed a quarter wavelength from the main waveguide, and an odd number of quarter wavelengths apart. When the transmitter is on, both tubes fire. This causes the waveguide sections, ATR to A and TR to B, to act as shorted quarter wave lines.

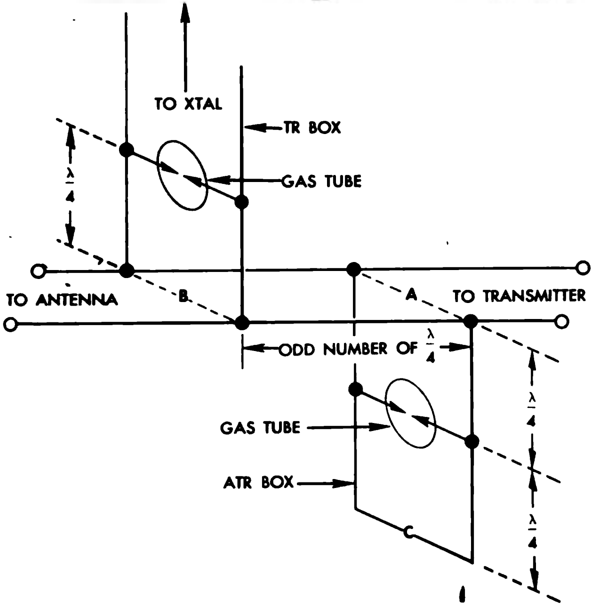
These sections then reflect open circuits at their junction to the main guide and the transmitter power is conducted to the antenna without loss. Since the TR tube fires, causing a short across that section, the transmitted power is prevented from entering the receiver.

When the transmitter is turned off, neither the TR nor the ATR tubes are fired. The section of waveguide the ATR tube is in now acts as a shorted half wave, since the distance between the end of the waveguide at C and point A is a half wave. This shorted half-wave section reflects a short at point A which in turn reflects an open at point B. The received signal therefore passes into the receiver.



TR and Anti-TR System





The Series Parallel Duplexer

Another type of duplexer commonly used is the series parallel type. This duplexer is used in the APS-23. Its action is very similar to the one shown before except that the ATR is placed in series with the transmitter while the TR is placed in parallel, as shown in the diagram on the next page.

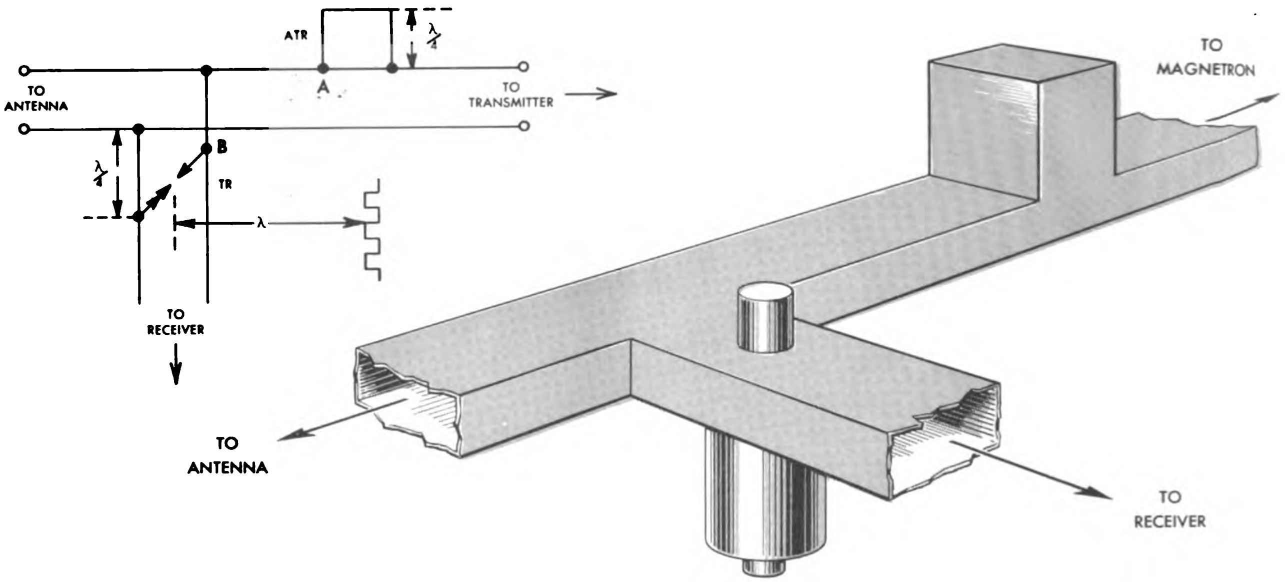
Notice that in this duplexer the ATR is placed in series with the magnetron. The TR tube is placed one-quarter wavelength from the junction of the main transmission line. When the transmitter fires, both the ATR and the TR are ionized. The impedance seen at the opening to the ATR box is a short circuit which effectively seals the opening and passes the energy on down the line. The impedance across the TR tube electrodes is also a short circuit, and at its junction with the main transmission line, reflects high impedance. The transmitter energy is then conducted on down the line to the antenna.

During reception, the ATR acts as a shorted quarter-wave stub reflecting a high impedance at point A. Since the distance from A to B is a full wave, the high impedance at point A is reflected as a high impedance at point B. The received signal therefore passes into the receiver.

TR Tubes

Unfortunately, a spark gap is not the ideal switch we have assumed. It behaves as a high impedance when not fired and as a nonlinear resistance that maintains a nearly constant low voltage across itself when fired. Its characteristics change somewhat with use. In order to increase its efficiency, the spark gap is enclosed in a partially evacuated glass envelope and connected to the transmission line through a resonant cavity, as shown in the illustration on page 407.

The purpose of the TR tube is to keep the RF energy out of the receiver. If this energy were allowed to get into the receiving channel of a radar set it would undoubtedly burn out the crystal mixer. To protect the crystal mixer, the TR tube must fire as soon as the RF pulse reaches it.



Series-Parallel Duplexer

To accomplish this, most TR tubes are filled with a mixture of argon gas and water vapor, and have a negative voltage applied to an electrode placed near one of the gap electrodes. This keep-alive electrode is shown in the illustration on page 407.

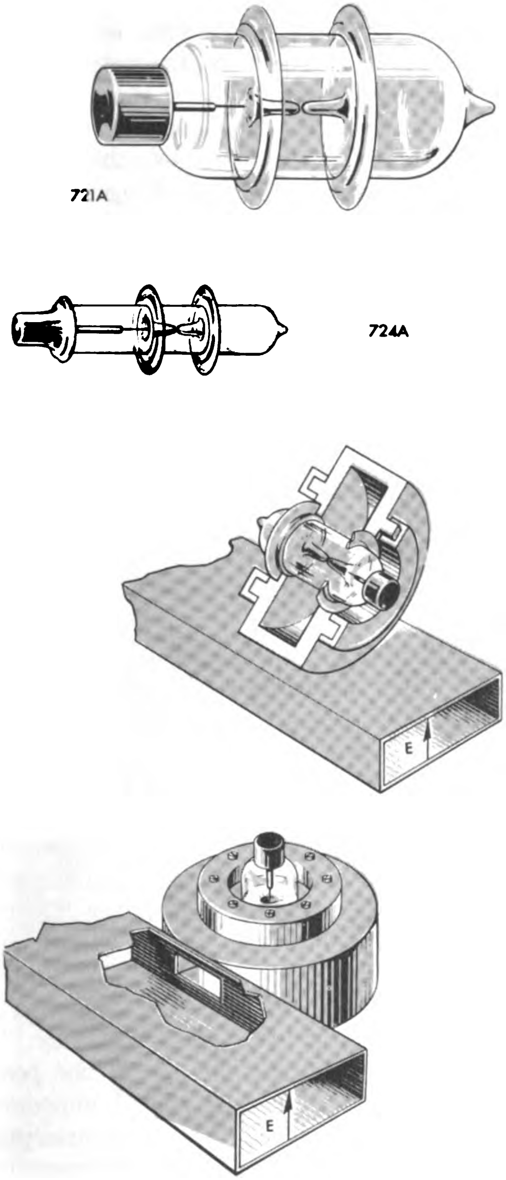
The negative voltage keeps a steady glow discharge near the enclosed spark gap. This discharge provides a continuous supply of ions and free electrons, and aids in establishing desirable conditions for an RF discharge path. When RF energy is applied to the electrodes, the tube fires more quickly.

Thus, the keep-alive electrode acts as a cathode while the back of one of the high frequency electrodes acts as the anode. In the modern airborne radar, the TR tube must fire within .01 microsecond after application of the transmitted pulse. Upon termination of the transmitted pulse, the tube must deionize at the end of 6 microseconds.

When a radar system is turned off, no keep-alive voltage is present. To avoid the possibility of strong pulses from near-by radars damaging the crystal mixer, some radar systems are provided with a crystal shutter or gate. Generally this consists of a metal vane which is automatically placed across the waveguide when the set is turned off.

After a glow discharge has been produced in a gas, and the voltage which produced that discharge has been removed, large numbers of free electrons and ions remain in the gas for a substantial time. The electrons and ions slowly recombine or migrate to electrodes in the surrounding structure.

Such ions are present in a TR tube following each transmitted pulse. Voltages resulting from received signals appear across the spark gap and set the ions in motion. A portion of the received power is therefore wasted in moving the ions. This means that the received signal is attenuated by the TR tube until recombination of all the ions has been accomplished. This is *recovery time,* and is defined as the interval between the transmitter pulse and the time at which the power in the signal applied to the receiver is one half as large as it would be if no ions were present. Normally the recovery time is approximately 6 microseconds. Recovery time is an important item to the observer since any increase in recovery time will add to the normal minimum range of his radar set. The theoretical minimum range of the APS-23 on the 30-nm range is approximately 184 feet, whereas, with a bad TR tube it may be as much as 8 to 10 miles. A scope photo showing a radar with a bad TR tube is shown in the illustration at the bottom of page 407.

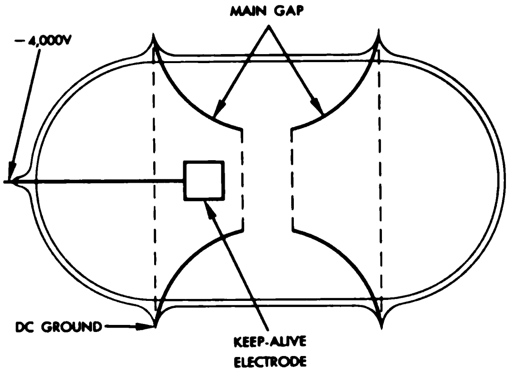


TR Tubes

ATR Devices

ATR tubes that are used in parallel with the transmitter are generally the same as the TR tube, except that no keep-alive voltage is applied. However, when the ATR device is used in series with the transmitter, it takes on a different form. The 1B35 ATR box is shown in the illustration on page 408.

Notice that the ATR box actually takes the place of a portion of the broad side of the waveguide. The ATR box consists of a cavity filled



TR Tube Showing Keep-Alive Electrode

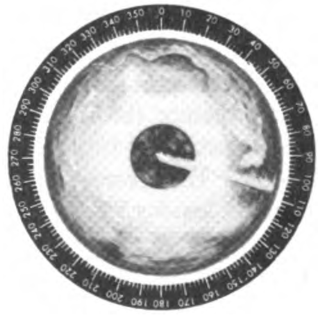
with gas (argon and water vapor). When the transmitter is on, the gas within the box ionizes and effectively completes the section of waveguide it replaces.

THE APS-23 DUPLEXER

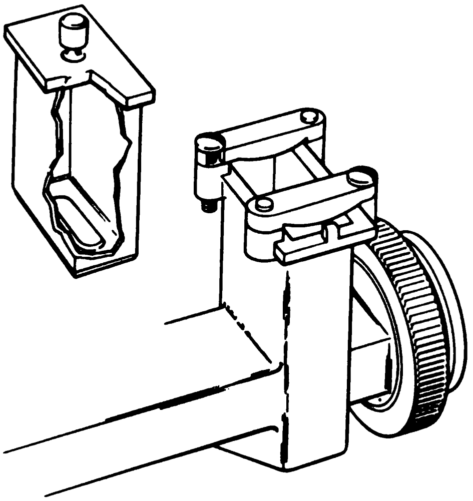
In your study of transmission lines, matching devices, and resonant cavities, perhaps you felt that it was all theory and very little of it was practical. To show how all of these various subjects tie together in a common piece of equipment, let us examine the APS-23 duplexer and mixer assembly which is shown in the illustration on page 409.

Notice that, at first glance, it appears to be several pieces of waveguide arranged in a random geometric pattern. However, let us start analyzing this assembly by first stating what it does. It has four major functions:

1. It is a duplexer and allows both transmitter and receiver to operate from one antenna.



Scope of Radar Set with Bad TR Tube



1B35 ATR Box and Mounting

2. It provides a mixing chamber where the radar received signal is heterodyned with the radar high frequency oscillator (HFO). The crystal diode, CR 1301, provides the non-linear element which allows the signals to mix. The IF frequency of 60 mc is taken from the crystal and sent to the receiver. If the receiver is to receive beacon signals, then the received beacon signal is heterodyned with the beacon HFO. The radar HFO is inoperative at this time.

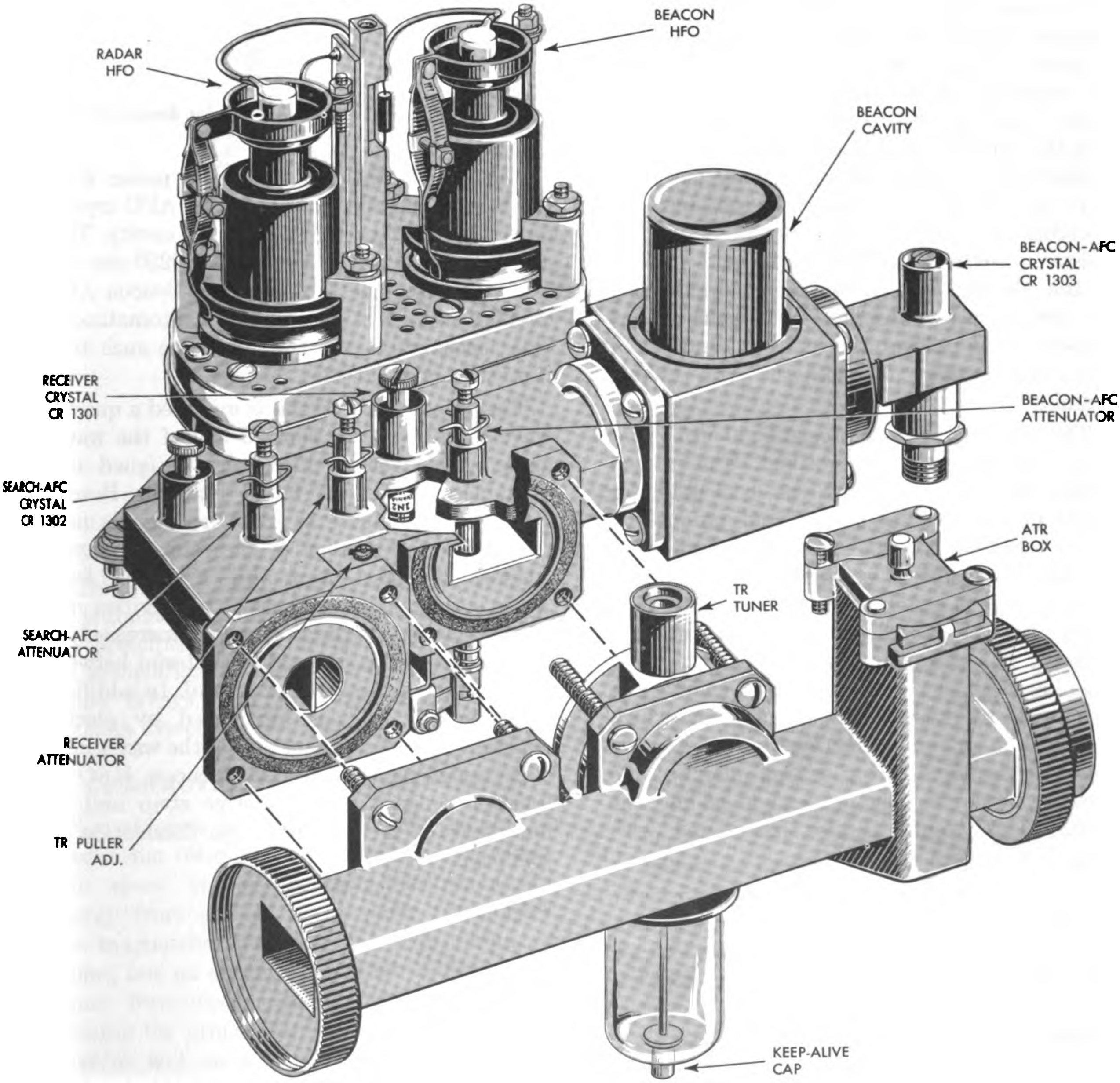
3. It provides a means by which a portion of the transmitted signal is detected by the Radar AFC Xtal, CR 1302. This signal is used by the Automatic Frequency Control Circuits to insure that the Radar HFO is always operating at the correct frequency to produce a 60-mc IF.

4. During beacon operation, the Beacon HFO oscillator operates continuously while the Radar HFO is inoperative. Since the received beacon signal (9310 mc) is different from the transmitting frequency (9375 mc), the set can no longer use a portion of the transmitted pulse for Beacon AFC. However, the duplexer contains a resonant cavity, tuned to 9250 mc, which acts as a reference to insure that the Beacon HFO is always tuned to 9250 mc. The Beacon AFC crystal, CR 1303, feeds this reference signal to the Beacon AFC circuits which in turn keep the Beacon HFO at the correct frequency.

Starting with the magnetron, let us discuss this assembly. Between the magnetron and the duplexer-mixer assembly is a short section of flexible waveguide. This section absorbs all dimensional tolerances and allows the magnetron and duplexer-mixer to be bolted together. Notice that choke joints are used between the magnetron and the flexible waveguide, and between the flexible waveguide and the duplexer-mixer to insure minimum loss of RF energy.

The 1B35 ATR box is placed in the broad side of the main waveguide, one wavelength from the entrance into the mixing chamber. The 1B24 TR tube is positioned a quarter wavelength from the junction of the main waveguide. Their combined action prevents the transmitted pulse from entering the receiver and channels the received signal into the mixing chamber.

The 1B24 TR tube is enclosed in a resonant cavity which makes its selectivity high, 50 mc between half-power points. This means that when the receiver is tuned for beacon operation, the loss in the received signal is such that only about one-hundredth of the received signal enters the mixing chamber. Loss is prevented by the TR puller which is inserted into the waveguide on the low power side of the TR tube, at a distance of half a wavelength from its out-put window. By adjusting the depth of the TR puller, the TR tube can be pulled into tune for the new (beacon) frequency. This puller is really nothing more than the post or tuning-screw method of matching impedances. The TR puller reflects a capacitive susceptance into the TR cavity causing it to resonate at a lower frequency. The receiver crystal (CR 1301) is placed one wavelength from the TR cavity so that it acts as a low impedance to the received signal and to the signal generated by the Radar HFO. The signal generated by the Radar HFO is coupled into the waveguide by the insertion of a probe as discussed on page 394. The probe is actually part of the HFO. The energy thus inserted into the waveguide is then coupled by means of matching irises into the Receiving and Radar AFC mixing chambers. Each window is tuned by an adjustment screw which acts as an attenuator and controls the amount of energy fed into the receiver and AFC mixing chamber. A strip of resistive material

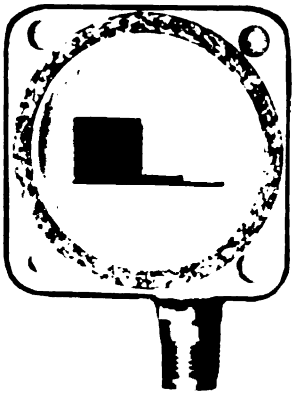


APS-23 Duplexer

terminates the waveguide behind the HFO to prevent reflections. A portion of the transmitted pulse is fed into the Radar AFC section. To avoid a large amount of power entering this section and damaging the crystal, this AFC section is joined to the main waveguide by means of a circular waveguide which is too small to allow the transmitted pulse to pass through unattenuated.

It is not practical to design a circular attenuator which will attenuate the signal the required amount because one must consider the reflections in the main waveguide and the mixer chamber due to the mismatch. Therefore, the circular attenuator is generally designed so that it attenuates the signal only partially. The rest of the attentuation is obtained by inserting a strip of carbon-coated bakelite resistance material so that its plane is in that of the electric field. The combined effect of the circular attenuator and the carbon-coated bakelite strip allows the correct amount of transmitted energy to enter the AFC mixing chamber.

The beacon HFO oscillates at 9250 mc so that radar beacons may be received. The beacon signal, 9310 mc mixes with the local oscillator signal in the receiver crystal and produces a 60-mc IF. The beacon HFO is mounted in another short section of waveguide. A portion of its energy is coupled through coupling irises to the receiver crystal and to the beacon AFC crystal. The iris coupling the beacon HFO chamber with the receiver mixing chamber is also tuned by a capacitive tuning screw which controls the amount of RF energy fed from the beacon chamber into the mixing chamber. To prevent reflections, the waveguide behind the beacon HFO probe is also terminated with a resistive strip.



Resistive Strip and Inductive Vane for Beacon HFO

The portion of RF energy that passes from the beacon chamber to the beacon AFC crystal must first pass through a resonant cavity. This cavity is pretuned to resonate at 9250 mc and acts as a reference to enable the Beacon AFC circuit to tune the Beacon HFO automatically. The specifications of this cavity are such that it must resonate at 9250 ± 1 mc.

The beacon AFC crystal is mounted a quarter wavelength from the shorted end of the waveguide. After the cavities were designed and manufactured, it was found that the Beacon HFO didn't oscillate at some frequencies near the resonant frequency of the resonant cavity when the beacon cavity fed into a matched load. To avoid having to redesign the cavities, the output loading of the cavity was increased by placing two strips of resistive material between the cavity and the detector crystal. In addition, an inductive iris was introduced by placing a metal vane out from one side of the waveguide. This metal vane allows the beacon HFO to function properly. The resistive strip and inductive vane are shown in the illustration on this page.

Chapter 37

Radiation and Antennas

This chapter deals with the basic principles of antennas. It discusses antennas in general, the principles of electromagnetic radiation and its application to radar antennas, various antenna arrays, and a typical airborne radar antenna system.

CHARACTERISTICS OF ANTENNAS

An antenna is an electronic device that is used either for radiating electromagnetic energy into space or for collecting electromagnetic energy from space. In the radar transmitter, the magnetron generates the high frequency signal, but an antenna is needed to change this signal into electromagnetic fields which are suitable for propagation into space. The radar receiver will amplify any signal that appears at its input terminals, but an antenna is required to intercept the electromagnetic fields that are in space and to change these fields into a voltage which the receiver can interpret.

Antenna Reciprocity

Fortunately, separate antennas are seldom required for transmitting and receiving radio energy, for any antenna transfers energy from space to its input terminals with the same efficiency with which it transfers energy from the output terminals into space — assuming, of course, that the frequency is the same. This property of interchangeability of the same antenna for transmitting and receiving operations is known as *antenna reciprocity*. Antenna reciprocity is possible chiefly because antenna characteristics are essentially the same regardless of whether an antenna is sending or receiving electromagnetic energy. Because of antenna reciprocity, most radar sets installed in aircraft employ the same antenna both for receiving and transmitting. An electronic switch in the radio frequency line first connects the single antenna to the transmitter, then to the receiver, depending upon the sequence of operation. Because of reciprocity of radar antennas, this chapter treats antennas from the viewpoint of the transmitting antenna with the understanding that the same principles apply equally well when the antennas are used for receiving electromagnetic energy.

Directional Properties

Usually, the most important characteristic of a radar antenna is its directional property, or simply its *directivity*. Directivity means that an antenna radiates more energy in one direction than in another. For that matter, all antennas are directional, some more than others. In radar operations, antennas are usually highly directive so that as much as possible of the electromagnetic energy generated by the transmitter will strike an object in a given direction. In other systems, it is desirable for the energy to be radiated equally well in all directions from the source. An example of an antenna system in radar which radiates energy in a given direction is the airborne navigation and bombing set. In this set, there is only a limited amount of power available at the transmitter. In order to achieve maximum benefit from this minimum power, all of it is sent in the same direction. Since the antenna in this set is also used for reception, it likewise receives electromagnetic energy from only one direction. Because of design features, it is possible to tell the direction of an object toward which this directional type antenna is sending energy and the direction from which the antenna is receiving energy. The physical position of the antenna indicates the direction of the object. An example of a nondirectional radar antenna is the antenna installed in the radar beacon. This antenna must receive energy equally well from all directions so that a radar-equipped airplane can ascertain its position regardless of its direction from the beacon antenna.

Antenna Efficiency

An efficient antenna is one which wastes very little energy. The higher the antenna efficiency, the less the energy loss. The greatest cause of decreased antenna efficiency during the conversion of RF energy into electromagnetic energy is the loss of energy in the form of heat.

Actually antenna efficiency may be high. The efficiency may be as great as 90% in a single antenna element. An entire antenna system, properly designed, can concentrate all its radiated energy in a given direction with very little being lost in other directions. A directional receiving antenna is more sensitive to signals in a given direction than a similar nondirectional antenna is to signals in all directions. This greater sensitivity, which is another way of stating that there is less energy loss by the antenna, represents greater receiving ability and, accordingly greater efficiency.

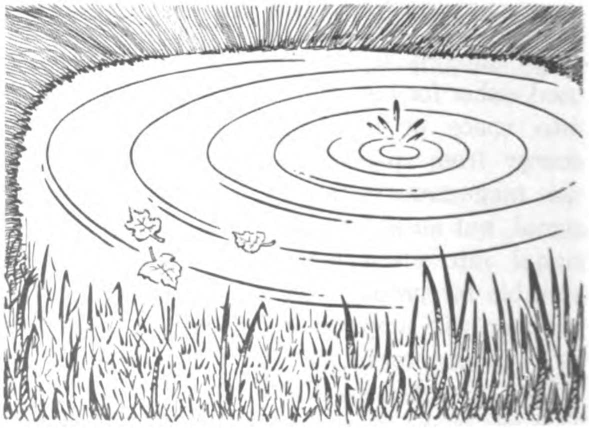
WAVE MOTION

The movement of waves in a pond and the movement of electromagnetic waves have some characteristics in common.

Wave Motion in Pond

Suppose you are standing at the edge of a small pond of water. The surface of the pond is smooth and unbroken except for a few dead leaves floating near your feet. If you toss a stone into the center of the pond, ripples spread out from the place where the stone struck the water. In a few moments, the disturbance reaches the leaves floating near you, and they begin to bob up and down. The energy imparted by the stone has made itself felt across the surface of the pond, as shown in the illustration below.

When the stone is tossed into the pond, it sets up a disturbance at the place where it strikes the water. With the disturbance as the center, waves are produced which move outward in expanding circles. As each circle moves out toward the edge of the pond, a new one



Wave Motion in Pond

forms at the center. Eventually, the waves cover the surface of the pond, and the leaves are bobbing up and down as each wave passes. Notice that the leaves are not carried along with the waves and washed up on the shore, but have only an up and down motion. This is because the water on which the leaves float is not moved sideways by the waves. The only effect of the wave is to raise or lower each particle in its path. The part of the wave that raises a particle is called the crest; the part that lowers a particle is called the *trough.*

As you noted, it took a certain amount of time for the waves to cover the distance from the disturbance to the leaves near the edge. This distance covered in a certain amount of time is a measure of the velocity or speed with which the waves travel across the pond. It can be expressed in terms of miles per hour, or centimeters per second, or any unit of distance per any unit of time. Velocity can also be expressed by the simple relationship,

$$\text{Velocity} = \frac{\text{distance traveled}}{\text{time required to travel this distance}}$$

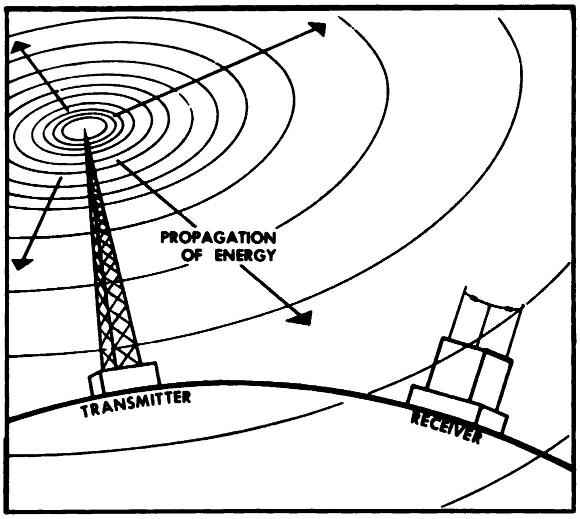
or,

$$v = \frac{d}{t}$$

Electromagnetic Wave Motion

Suppose that a source of RF energy, such as a radio transmitter, is connected to an antenna. The antenna sets up fields in the surrounding space, and the resultant waves are propagated in all directions in the form of expanding concentric circles, as shown in the illustration to the right. When these waves encounter some device, such as a receiving antenna, on which they can act, some of their energy is given up to that device. The essential fact in this picture is that energy is transmitted by means of wave motion — in this case by electromagnetic waves. Some of the energy produced by the transmitter is transferred through space to the receiver in somewhat the same way as some of the energy of the stone tossed in the pond was transferred through the water.

The velocity of these electromagnetic waves moving through space is found in the same manner as the velocity of the water waves. It is the distance divided by the time required to cover that distance. Instead of taking any distance, suppose you consider the distance from the crest of one wave to the crest of the next. This is the length of one wave, or one *wavelength.* The same distance can also be measured from the trough of one wave to the trough of an adjacent wave. The electromagnetic wave in space has definite wavelength which can be measured in the manner just described. It is usually measured in meters, and its symbol is the Greek letter lambda (λ).



Electromagnetic Wave Motion in Space

The time required for an electromagnetic wave to travel a distance equal to its wavelength is called the *period* of the wave. The symbol for this time is the letter T. Going back to the basic relationship given previously, you recall that velocity equals distance divided by time, or

$$V = \frac{\lambda}{T}$$

where V is the velocity of the electromagnetic wave in meters per second, λ is the wavelength in meters, and T is the period in seconds.

If there are a million complete waves generated every second, then the time required for one wave to travel a distance of one wavelength is one-millionth of a second. Now, the number of complete waves that occurs every second is referred to as the frequency (f) of the wave. The frequency is therefore a reciprocal of the period, or

$$T = \frac{1}{f}$$

Substituting in the previous equation,

$$V=\frac{\lambda}{1/f}$$

or,

$$V=f\lambda$$

The actual velocity of an electromagnetic radio wave in space has been found to a high degree of accuracy. The exact figure is slightly less than 300,000,000 meters per second. This is equivalent to 186,000 statute miles per second, or 984 feet per microsecond. This is also the velocity of light, another form of electromagnetic radiation. The velocity of electromagnetic energy through space is constant. So suppose you rearrange the above equation to make it more useful as follows,

$$\lambda=\frac{V}{f}$$

and,

$$f=\frac{V}{\lambda}$$

These are important basic equations dealing with wave motion. When you use them, it is important that you employ consistent units. Commonly, V is expressed in terms of meters per second, f is expressed in cycles, and λ is in meters.

Examples: What is the wavelength corresponding to a radio frequency of 5 mc (5,000,000 cycles per second)? When you use the first of the two equations given above,

$$\lambda=\frac{V}{f}$$

$$=\frac{300{,}000{,}000}{5{,}000{,}000}$$

$$=60 \text{ meters}$$

What is the frequency which corresponds to a wavelength of 10 meters? Using the second of the two equations given above,

$$f=\frac{V}{\lambda}$$

$$=\frac{300{,}000{,}000}{10}$$

$$=30{,}000{,}000 \text{ cps}$$

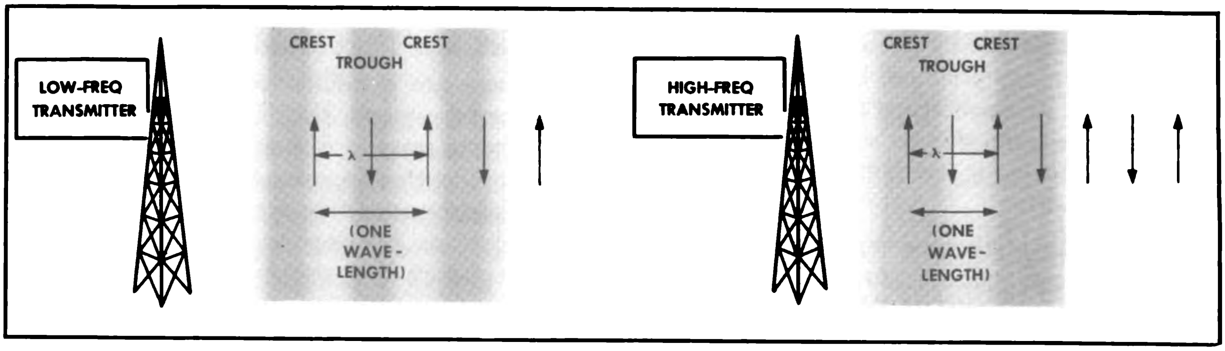
$$=30 \text{ mc}$$

Note from the equations and from the two examples given that wavelength and frequency are inversely related. As the frequency is increased, the wavelength is shortened; as the frequency is decreased, the wavelength is lengthened. This inverse relationship is shown in the illustration below. It is common to refer to the higher radio frequencies as short waves and to the lower radio frequencies as long waves.

Since an electromagnetic wave travels on an antenna, the antenna, too, is measured in wavelengths. Because of the resistance of the wire, the movement of waves along a wire (antenna) is somewhat slower than in space. Wavelength on a wire, therefore, is slightly less than that of a wave traveling in space and is expressed by the equation:

$$\lambda a=\frac{3\times10^{8}\times.94}{f} \text{ meters}$$

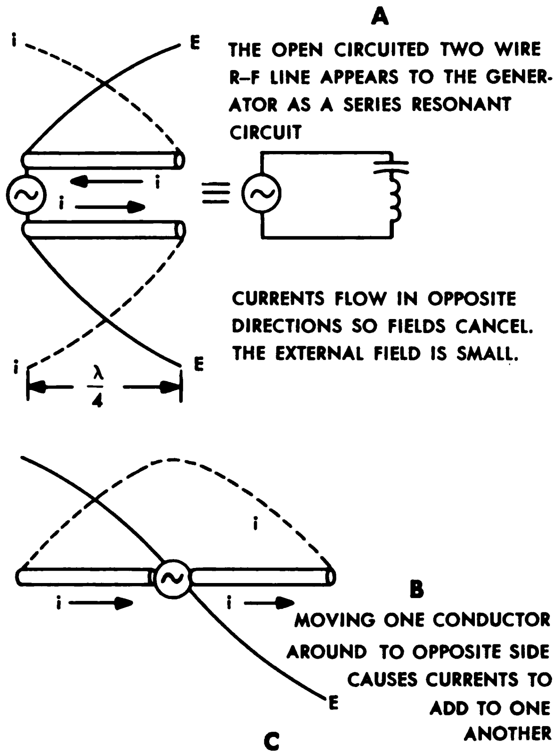
Physically, a half-wave antenna is about 6% shorter than a half wave traveling in space. (This accounts for the fact that it is necessary to multiply the wavelength in the antenna formula by .94.) In this manual, assume that the correction in antenna length has been made whenever the length of the antenna is given in wavelength.



Relationship of Wavelength and Frequency

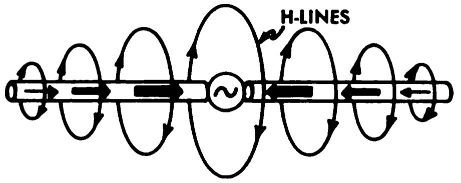
ELECTROMAGNETIC ENERGY IN SPACE

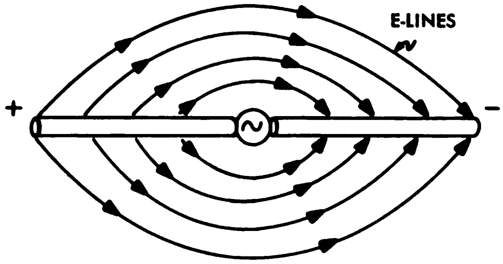
Because a half-wave dipole is the fundamental element in an antenna system, it can be used as a starting point for discussing the radiation of electromagnetic energy into space. Electrically, you can think of the dipole as an open-circuited quarter-wavelength RF line that is excited by a generator and appears to the generator as a resonant circuit. In diagram A of the illustration to the right, a dipole consisting of a piece of wire cut in half is attached to the terminals of a high-frequency AC generator. The frequency of the generator is such that each half of the wire is one-quarter the wavelength of the generator output. Due to the resonant characteristics of this arrangement, high circulating currents will flow in the antenna when it is excited by the generator. This current flow sets up two strong fields about the antenna—one the magnetic field (H) and the other the electric field (E). However, due to the U-shape of the antenna, and to the fact that the currents in each wire move in opposite directions, the H-fields and E-fields are in opposition and cancel, making the overall field strength of the antenna assembly very low. Suppose, however, that the two wires are moved as far apart as possible, as shown at B. In this case, the current flow in each wire is in the same direction, and the resulting E-fields and H-fields no longer oppose, but aid each other and create strong fields about the antenna. You can see the distribution of E-field and H-field in diagrams C and D. Later you will see that one of these fields is responsible for electromagnetic radiation.


A
i
E
THE OPEN CIRCUITED TWO WIRE R–F LINE APPEARS TO THE GENERATOR AS A SERIES RESONANT CIRCUIT
i
i
≡
CURRENTS FLOW IN OPPOSITE DIRECTIONS SO FIELDS CANCEL. THE EXTERNAL FIELD IS SMALL.
i
E
λ/4
i
i
i
B
MOVING ONE CONDUCTOR AROUND TO OPPOSITE SIDE CAUSES CURRENTS TO ADD TO ONE ANOTHER
E

C

DISTRIBUTION OF H–FIELD AROUND HALF WAVE ANTENNA





D

DISTRIBUTION OF E–FIELD AROUND HALF WAVE ANTENNA. (ONE QUARTER CYCLE AFTER C)

Development of Half-Wave Antenna

The distribution of current and voltage about an antenna is also shown at B in the illustration. The dotted line represents the current distribution. Note that most of the current flows at the center and not at the ends of the antenna. The voltage, on the other hand, is maximum at the ends and minimum at the center. The magnetic field is greatest where the current is greatest, as you can see at C. D shows the E-field when the voltage difference between the ends of the conductor is greatest. Note that about the antenna as in the case of a resonant circuit, where voltage is maximum, the E-field is maximum and the H-field is minimum. Both of these fields vary at a sinusoidal rate with a time difference of one-fourth cycle or 90° between them.

Fields associated with energy stored in antennas are called *induction* fields. The strength of these fields decreases with the square of the distance from the antenna. Induction fields are local in effect and play no part in the transmission of electromagnetic energy. They represent only the stored energy in the antenna and are responsible only for the resonant effects the antenna reflects to the generator. The fields set up in the transfer of energy through space are known as the *radiation* fields. Although these fields decrease with the distance from the antenna, this decrease is much less rapid than the decrease of the induction field. The decrease of the radiation field is linear and not according to the square of the distance. Therefore, the radiation fields reach great distances from the antenna. The radiation fields are responsible for electromagnetic radiation into space.

Radiation of Fields

The exact cause and the mechanism for the phenomenon of radiation is not known to the electronic scientist. However, you can get a pretty good idea of what takes place by studying the illustration at the top of page 417. It is a simple picture of an E-field detaching itself from an antenna. At A, the E-field is maximum. Notice that the outer E-lines are stretched away from the inner ones. This is due to the repelling force between the lines of force. As the voltage decreases toward zero, the radiation field decreases and the E-lines retract back into the conductor. If the voltage decreases slowly, the entire field will collapse back into the conductor. On the other hand, if the decrease of voltage is rapid, the outer parts of the field cannot move in very fast, and when the voltage is almost zero, as shown at B, a relatively large E-field will still exist around the conductor. At C an exact zero voltage condition is shown. The E-field in this case does not reach zero, and it is left with no voltage to support it. When the next E-field develops around the antenna, it will push away the preceding E-field in the manner shown at D. The action of one field pushing away the preceding field is analogous to the snapping of a whip, in that part of the E-field is snapped off the antenna with each cycle. The snapped-off field is projected into space and moves away from the antenna at a constant speed of 186,000 miles per second. A similar action projects H-fields into space.

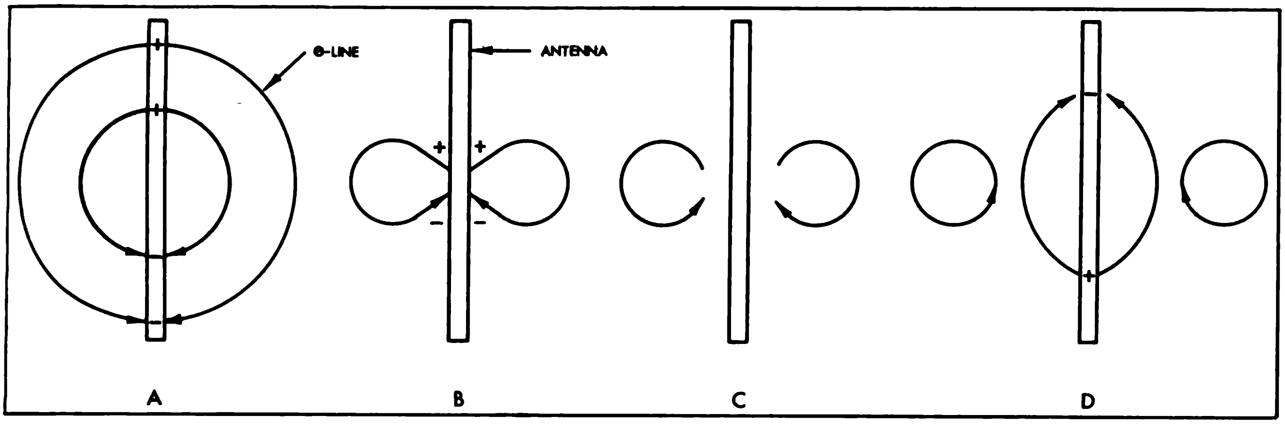
Another factor to remember about radiation is that the ease with which it may occur varies with frequency. At lower frequencies, such as 60 cps power frequency, for example, voltage on the antenna changes so slowly that the component of energy radiated is extremely small and is of no practical value. At higher frequencies (above 50,000 cps), the radiated energy is great enough to meet communications requirements.

Fields in Space

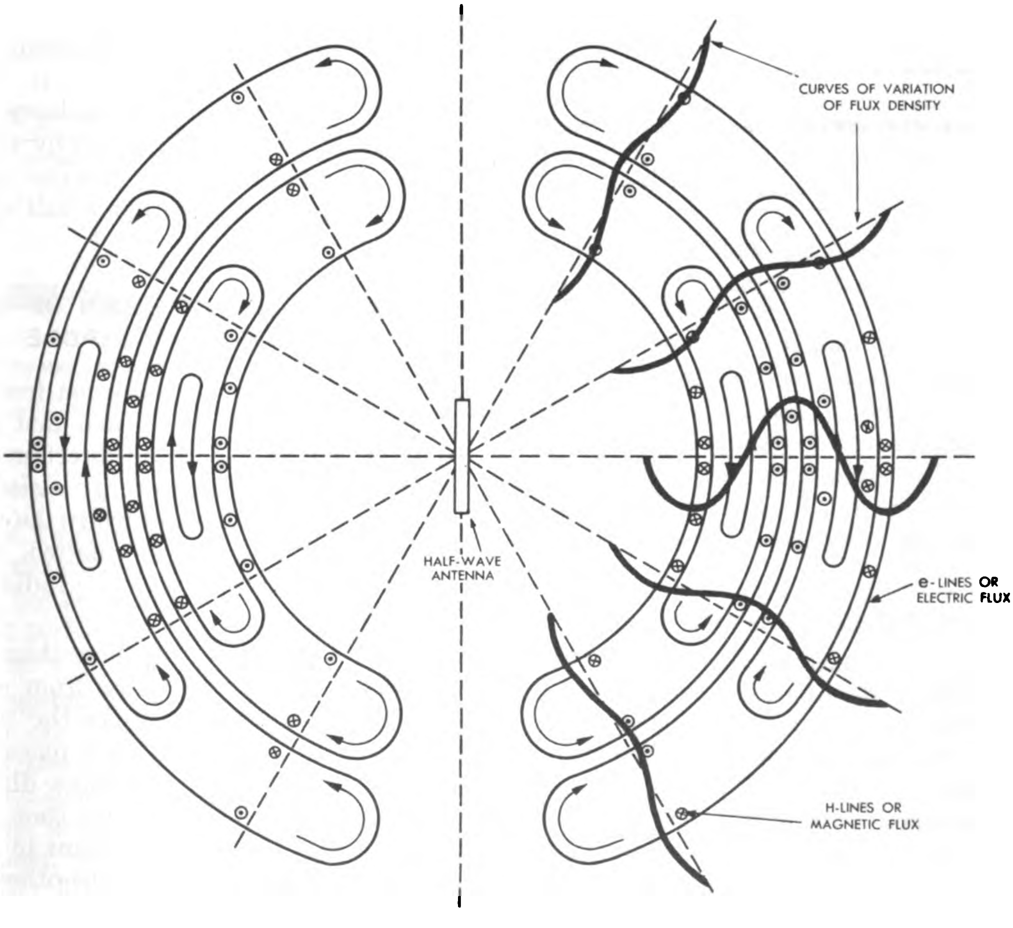
The lower illustration on page 417 shows the manner in which radiation fields are propagated from the antenna. The E-field and the H-field are shown as separate sets of flux lines about the antenna. The electric flux lines are the closed loops on each side of the antenna. The magnetic flux lines are closed circular loops that have their axes around the antenna, or in other words, the antenna is the center of each loop. They are represented as dots and crosses. The sine waves, labeled Curves of Variation of Flux Density, indicate the relative field strength at various distances and angles from the antenna. Since the field configuration is not a standing wave but a traveling wave phenomenon, the magnetic and electric fields are in phase; therefore, the sine waves apply to the flux density of either field.

In the directions perpendicular to the axis of the antenna, both the H-field and the E-field are strong, for this is the direction in which both fields originally were formed. Parallel to the axis of the antenna, or off the ends, there is no H-field, and only a very small E-field is formed. The flux density, therefore, is small in these two directions. Due to the directional properties of the half-wave antenna, most of the radiated energy travels in directions perpendicular to the antenna, and very little in the directions along its axis.

As previously stated in the discussion of waveguides, no moving E-field can exist without an H-field; conversely, no H-field can be propagated without an associated E-field. Similarly, with the propagation of electromagnetic energy into space, no moving electrostatic forces can exist without magnetic stresses existing in space; no magnetic force in motion can exist without an electrostatic stress. Each creates the other and one cannot exist without the other. The direction of motion is the direction in which the fields travel away from the antenna. The current associated with the magnetic field does not flow since in space there is no conductor to carry current; however, the



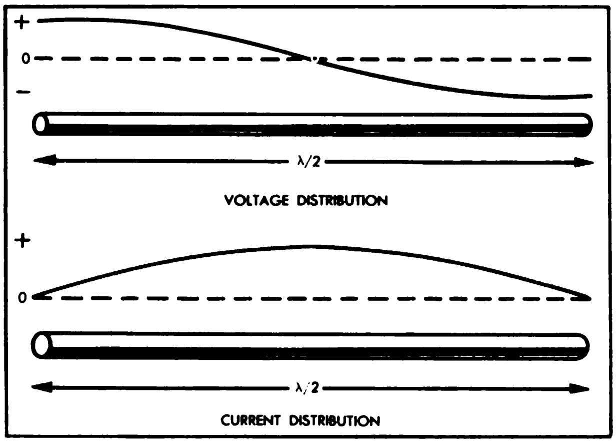
E-field Detaching Itself from an Antenna



Fields in Space around a Half-wave Antenna field exists. To visualize a field existing without current flow, think of a moving magnetic field cutting a glass rod. A voltage (electrostatic stress) is induced in the rod, but there is little actual electron movement because the rod is a good insulator. Magnetic lines move in space and set up electric stresses in space in a similar manner.

THE HALF-WAVE ANTENNA

As stated previously in the discussion of electromagnetic radiation, a half-wavelength conductor is the simplest of the radiating elements. Considerable radiation occurs in this element because of its resonant characteristics and its ability to store large amounts of energy in its induction fields. Resonance causes high voltages and high circulating currents and they, in turn, produce strong fields around the antenna.



Current and Voltage Distribution of a Half-wave Dipole

Current and Voltage Distribution

According to the accompanying illustration showing voltage distribution in a half-wave antenna, the voltage standing wave is high at the ends of the antenna and low at the center. As previously explained, this is also the case with the quarter-wave open circuit two-wire line from which the half-wave antenna is developed. An examination of the current distribution curve shows that the current standing wave reaches maximum a quarter cycle after the voltage. The current is maximum at the center and zero at the ends where there is no place for electrons to go. In practical applications, the ends of a half-wave antenna must be insulated due to the high voltages there, and the center of the antenna must have low resistance in order to minimize the I^2R losses due to the high current there.

Effect of Length on Antenna Impedance

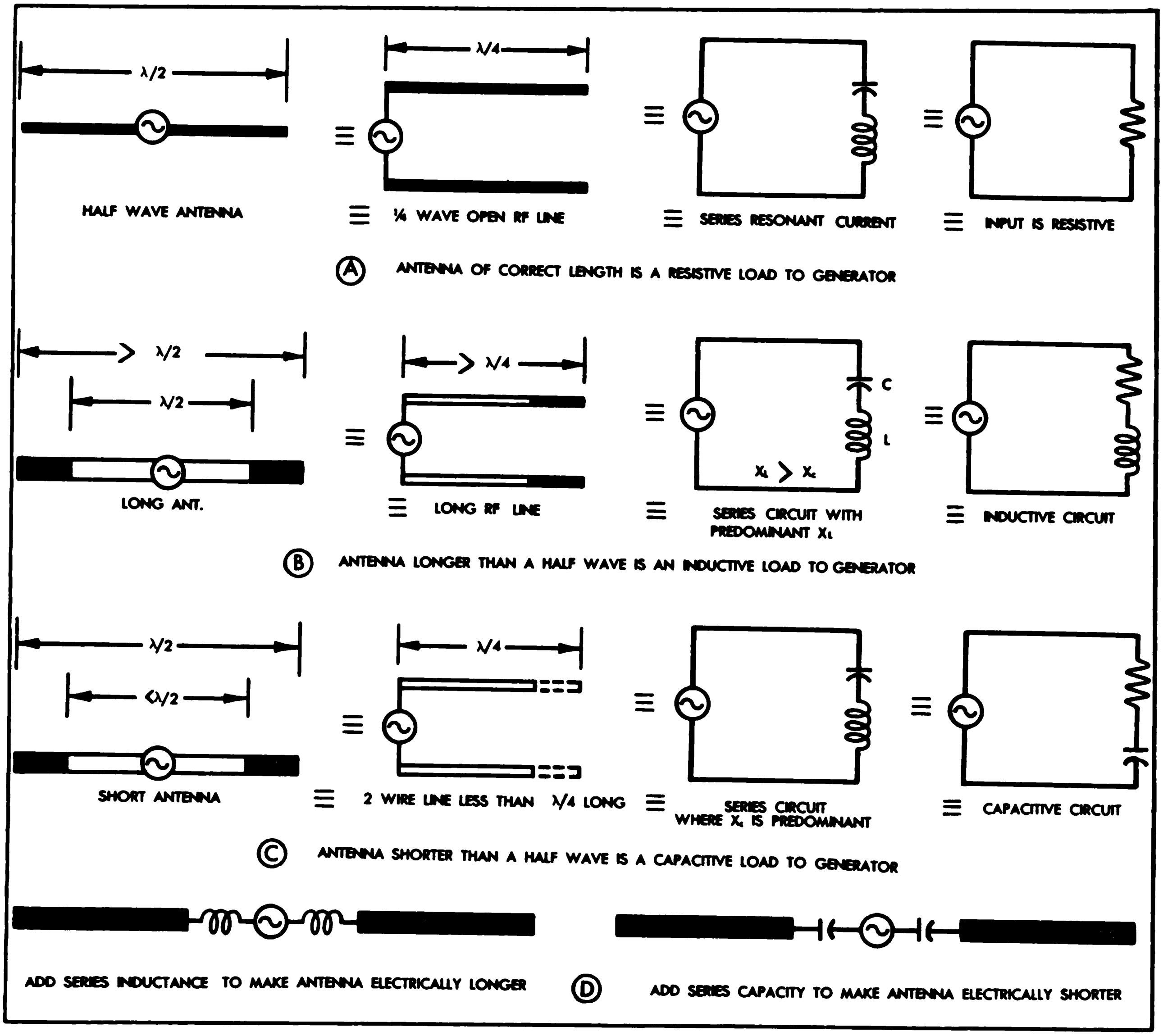
An antenna of the correct length acts like a resonant circuit and presents pure resistance to the excitation circuit as shown at A in the illustration on the next page. An antenna having other than the correct length displays both resistance and reactance to the excitation circuit. An antenna slightly longer than a half wave, for example, acts like an inductive circuit. This is understandable if you think of the antenna in terms of a quarter wave RF line. When the antenna is excited at the center, it is equivalent to a $\lambda/4$ RF line, looking at it from the generator end. Any two-wire line, longer than a quarter wave, is like a quarter wave section with an open-circuited section added to it. The open section, which is capacitive, is inverted to inductance at the input terminals by the quarter wave. In the same manner, you can see that an antenna slightly shorter than a half wave looks capacitive to the input.

RADIATION PATTERN OF THE HALF-WAVE DIPOLE

If you measured the radiation from a single point in space, you would find that this point radiates equally well in all directions, and the strength of the radiated energy varies inversely as the distance. If all points where the energy was of the same strength were plotted, the points would form a sphere, with the radiating point at the center.

If radiation occurs from more than one point in space, the radiated signal from each point will add vectorially to produce the total radiation strength. Radiation takes place from two different parts of the half wave dipole. It is, therefore, reasonable to assume that the energy from one part may, at some point in space, add to, or cancel the energy from the other part.

After a vector addition has been made of the energy at the infinite points surrounding a half-wave dipole, all of the points of equal strength are plotted. It is found that the resultant dia-


λ/2
HALF WAVE ANTENNA
λ/4
≡ ¼ WAVE OPEN RF LINE
≡ SERIES RESONANT CURRENT
≡ INPUT IS RESISTIVE
Ⓐ ANTENNA OF CORRECT LENGTH IS A RESISTIVE LOAD TO GENERATOR
> λ/2
λ/2
LONG ANT.
> λ/4
≡ LONG RF LINE
C
L
$X_L > X_C$
≡ SERIES CIRCUIT WITH PREDOMINANT X_L
≡ INDUCTIVE CIRCUIT
Ⓑ ANTENNA LONGER THAN A HALF WAVE IS AN INDUCTIVE LOAD TO GENERATOR
λ/2
<λ/2
SHORT ANTENNA
λ/4
≡ 2 WIRE LINE LESS THAN λ/4 LONG
≡ SERIES CIRCUIT WHERE X_C IS PREDOMINANT
≡ CAPACITIVE CIRCUIT
Ⓒ ANTENNA SHORTER THAN A HALF WAVE IS A CAPACITIVE LOAD TO GENERATOR
ADD SERIES INDUCTANCE TO MAKE ANTENNA ELECTRICALLY LONGER
Ⓓ
ADD SERIES CAPACITY TO MAKE ANTENNA ELECTRICALLY SHORTER

Effect of Length of Antenna on Impedance

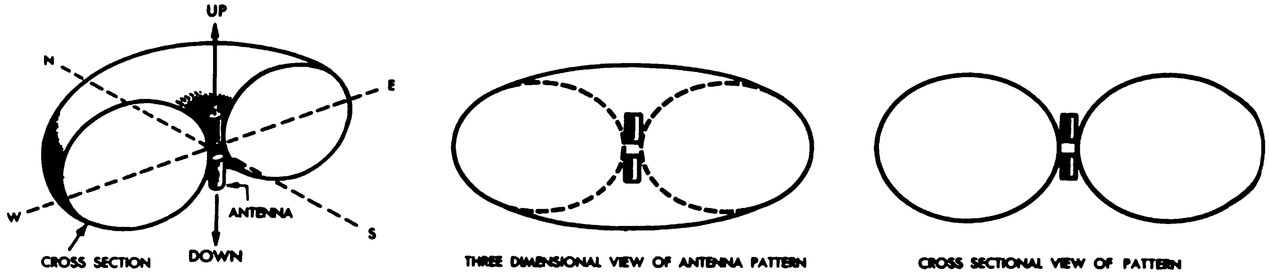
gram is not a sphere, but is a doughnut-shaped, three-dimensional figure as shown in the illustration on the next page. Note that it resembles two almost perfect circles which are adjacent to one another.

Polar diagrams of a half-wave dipole antenna are shown on page 421. These indicate the strength of the radiation field. The strength of the radiation field is called field strength and is measured in units called volts-per-meter. One unit is defined as the field strength which will produce one volt in a conductor one meter long. As field strengths of a volt-per-meter are seldom encountered, weaker fields are measured in microvolts or millivolts per meter.

IMPEDANCE AT INPUT TERMINALS

In practice an antenna acts like a resonant or near resonant circuit and, theoretically, like a perfectly tuned circuit. When it looks like a series resonant circuit, its input terminals must display zero impedance, and when it is connected like a parallel resonant circuit, it must display infinite impedance. Further, if it is not of the correct length, it must display reactance as well as resistance.

Practical antennas, like practical resonant circuits, have losses which must be replaced by the source. These losses make the input impedance of the antenna somewhat resistive.



CROSS SECTION

THREE DIMENSIONAL VIEW OF ANTENNA PATTERN

CROSS SECTIONAL VIEW OF PATTERN

Radiation Pattern of Half-wave Antenna

This resistance is a combination of two resistances — the resistance of the conductors themselves, which is increased by the skin effect at high frequencies, and the radiation resistance.

Radiation resistance is a fictitious resistance which dissipates the same amount of power in the form of heat that is actually dissipated as radiated energy. Because of radiation resistance, an antenna allows part of its fields to escape into space. This makes it necessary for additional current to flow into the antenna from the source of power to replace the part of the field that escaped. Therefore, the source must provide the energy lost by radiation. The energy lost is power (P), and its ratio to the square of the current (I) is the radiation resistance, R_r. Mathematically, it is expressed,

$$R_r = \frac{P}{I^2}$$

Radiation resistance may therefore be defined as the ratio of the power radiated to the square of the current in the antenna. Radiation resistance and conductor resistance constitute the total input resistance to an antenna. It is possible to make the conductor resistance small in comparison to the radiation resistance by using large diameter, low resistance conductors. It is desirable that a major portion of the input resistance become the radiation resistance. In practice, radiation ratios as high as 9:1 are obtainable. This means than an antenna can be 90% efficient as a radiator.

A graph of the input resistance at the center of antennas of various lengths is shown on the next page. For the half-wave dipole $(L=\lambda/4)$, the input is 73 ohms. The input resistance of a full wave antenna $(L=\lambda/2)$ is as high as 9000 ohms, depending on the diameter of the conductor. For the lengths of antennas in which the reactance is zero, the input resistance is the input impedance of the antenna.

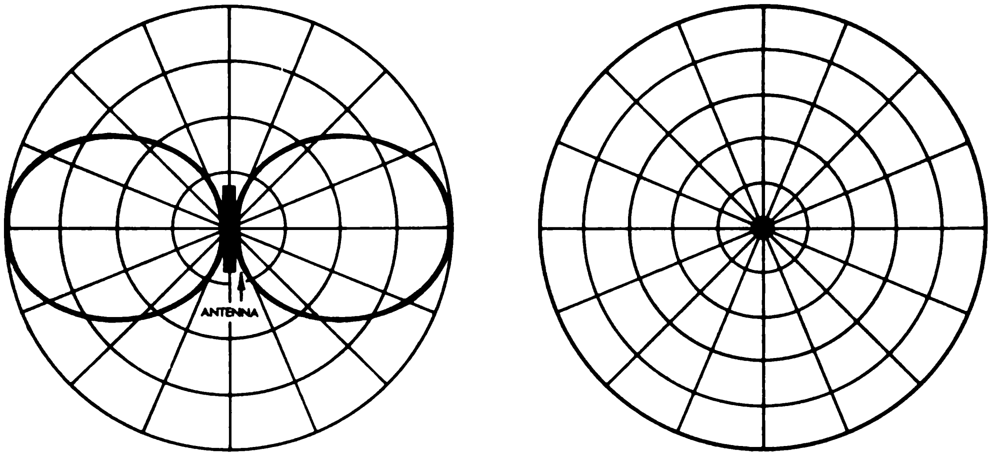
Factors which affect the input impedance of an antenna are nearby conducting objects, such as the earth, and other antennas or the skin of the aircraft. The graph applies only to a center-fed antenna which is located in free space and not close to any conductor.

POLARIZATION OF AN ELECTROMAGNETIC WAVE

Electromagnetic fields in space are said to be *polarized* and the direction of the electric field is considered the direction of polarization. As the electric field is parallel to the axis of a half-wave dipole, the antenna is in the plane of polarization. When the half-wave dipole is horizontally orientated, the emitted wave is horizontally polarized. A vertically polarized wave is emitted when the antenna is erected vertically. Horizontal polarization and vertical polarization are shown in the illustration on page 422.

For maximum absorption of energy from the electromagnetic fields, a half-wave dipole should be located in the plane of polarization. This places the conductor at right angles to the magnetic lines of force and parallel to the electric lines.

In general, the polarization of a wave does not change over short distances. Therefore, transmitting and receiving antennas are orientated alike, especially if short distances separate them. Over long distances, the polarization changes. The change is usually small at low frequencies. At high frequencies, the change is quite rapid.



POLAR DIAGRAM FOR HALF WAVE ANTENNA IN ANY PLANE CONTAINING THE ANTENNA

POLAR DIAGRAM FOR HALF WAVE ANTENNA IN PLANE PERPENDICULAR TO ANTENNA

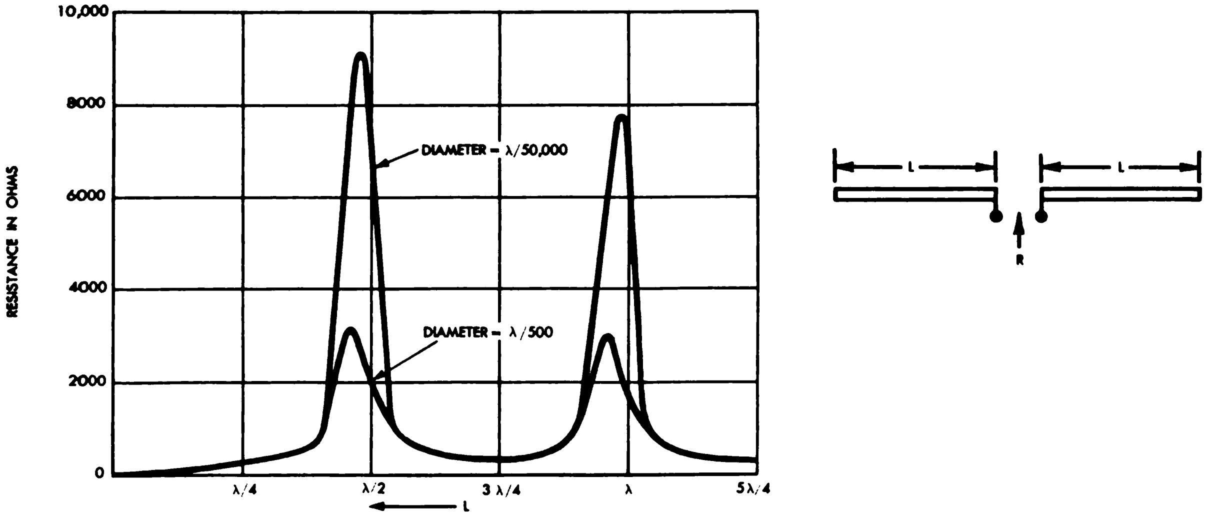
Polar Diagram of a Half-wave Dipole

With radar transmission, a received signal is a reflected wave from some object. As the polarization of the reflected signal varies with the type of object, no set position of the receiving antenna is correct for all returning signals. Generally, the receiving antenna is polarized in the same direction as the transmitting antenna.

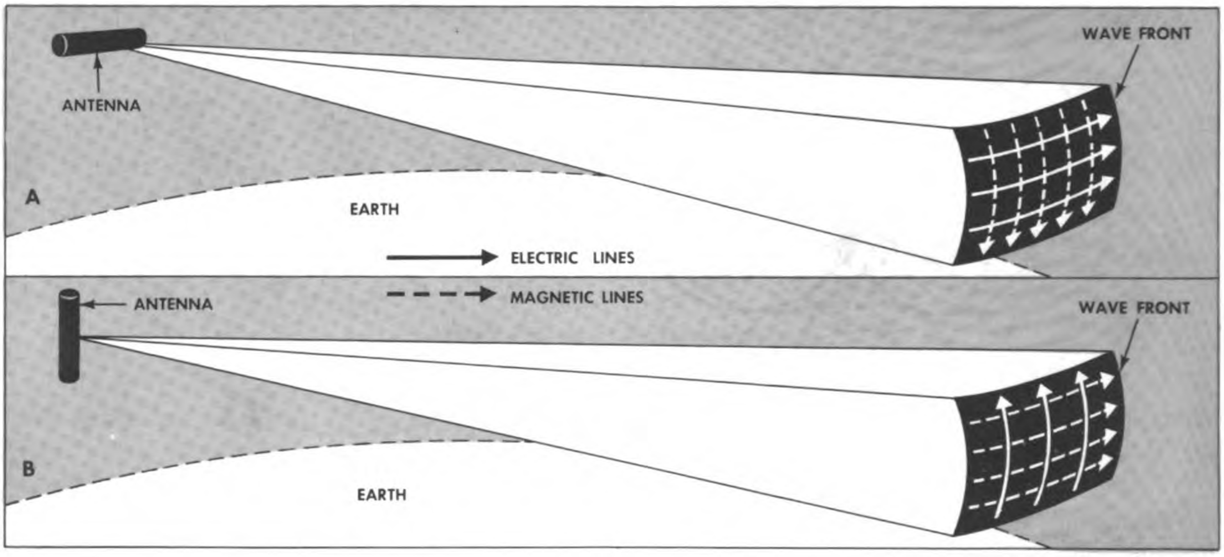
When the transmitting antenna is close to the ground insofar as propagation is concerned, vertically polarized waves cause a greater signal strength along the earth's surface. On the other hand, antennas high above the ground should be horizontally polarized to get the greatest signal strength possible. For this reason most airborne radar systems radiate horizontally polarized waves.

INTRODUCING ENERGY TO THE ANTENNA

Although energy may be fed to an antenna in a variety of ways, most antennas are *voltage fed* or *current fed* as shown in the illustration at the bottom of the next page. When the excitation energy is introduced to the antenna at the point of high circulating currents, the antenna is current fed. When the energy is introduced at the point of maximum voltage, the antenna is voltage fed.



Resistance at Input of Center-Fed Antenna

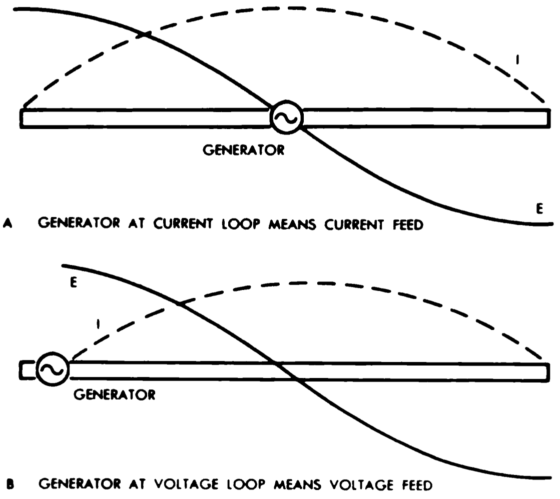


Horizontal and Vertical Polarization

It is seldom possible to connect a generator directly to an antenna. Usually RF transmission lines carry the energy from the generator to the antenna. The RF lines which carry the excitation energy might be resonant lines, nonresonant lines, or a combination of both.

Feeding Antennas with Resonant lines

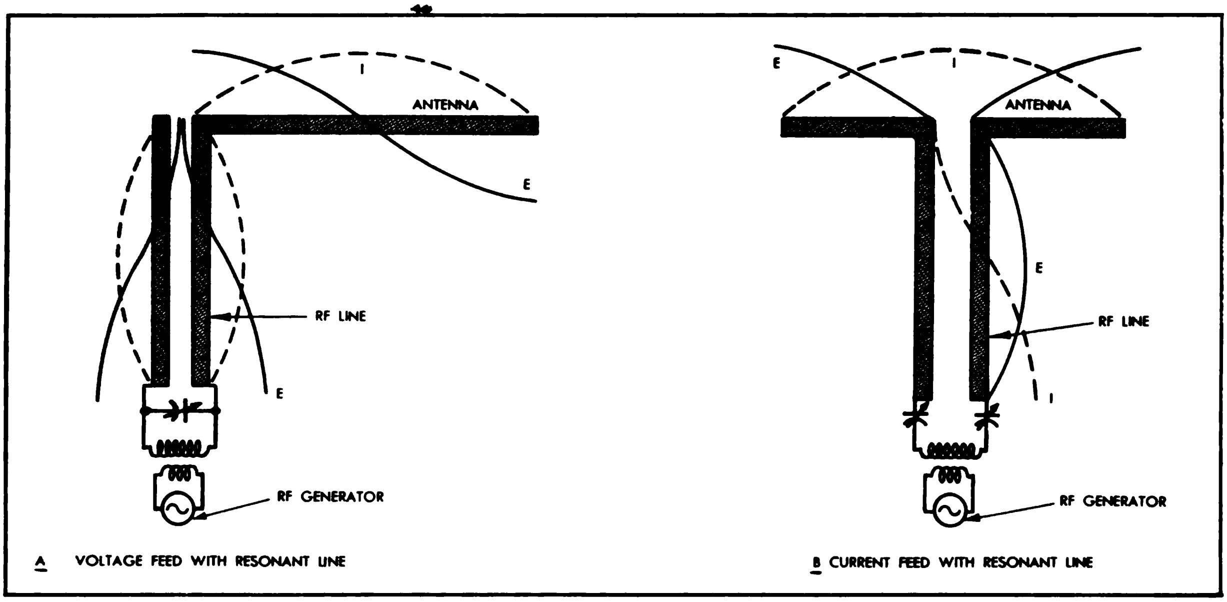
In a voltage-fed half-wave antenna which is excited by a resonant RF line, the end of the antenna is connected to one side of the line, as shown at A in the illustration on the next page. Voltage changes at that point excite the antenna into oscillation. The impedance at the end of the antenna is high since the voltage is high and the current is low there. The RF line shown is a half-wavelength, two-wire line. The characteristic impedance of such a line is usually low in order that the high impedance at which it is terminated will set up standing waves on it. In addition it is always a multiple of quarter wavelengths, electrically, so that it will act like a resonant circuit. This makes the input to the line also a high impedance. A parallel resonant circuit is used to develop a high voltage across this high input impedance. Small irregularities in line lengths can be compensated for by tuning the parallel resonant circuit at the input.


I
GENERATOR
E
A GENERATOR AT CURRENT LOOP MEANS CURRENT FEED
E
I
GENERATOR
B GENERATOR AT VOLTAGE LOOP MEANS VOLTAGE FEED

Current and Voltage Feed

In the current-fed antenna with a resonant line shown at B, the RF line is connected to the center of the antenna. This antenna has a very low impedance at the center and, like the voltage-fed antenna, has standing waves on it. It is also a multiple of quarter wavelengths and is a resonant circuit. Making it exactly a half-wavelength causes the impedance at the sending end to be low. A series resonant circuit is therefore used to develop the high current necessary to excite the line. Adjusting the condensers at the input compensates for irregularities in line and in antenna length.

Although these examples of feeding antennas are simple ones, the principles described apply


I
ANTENNA
E
RF LINE
E
RF GENERATOR
A VOLTAGE FEED WITH RESONANT LINE
E
I
ANTENNA
E
RF LINE
I
RF GENERATOR
B CURRENT FEED WITH RESONANT LINE

Feeding Antennas with Resonant Lines

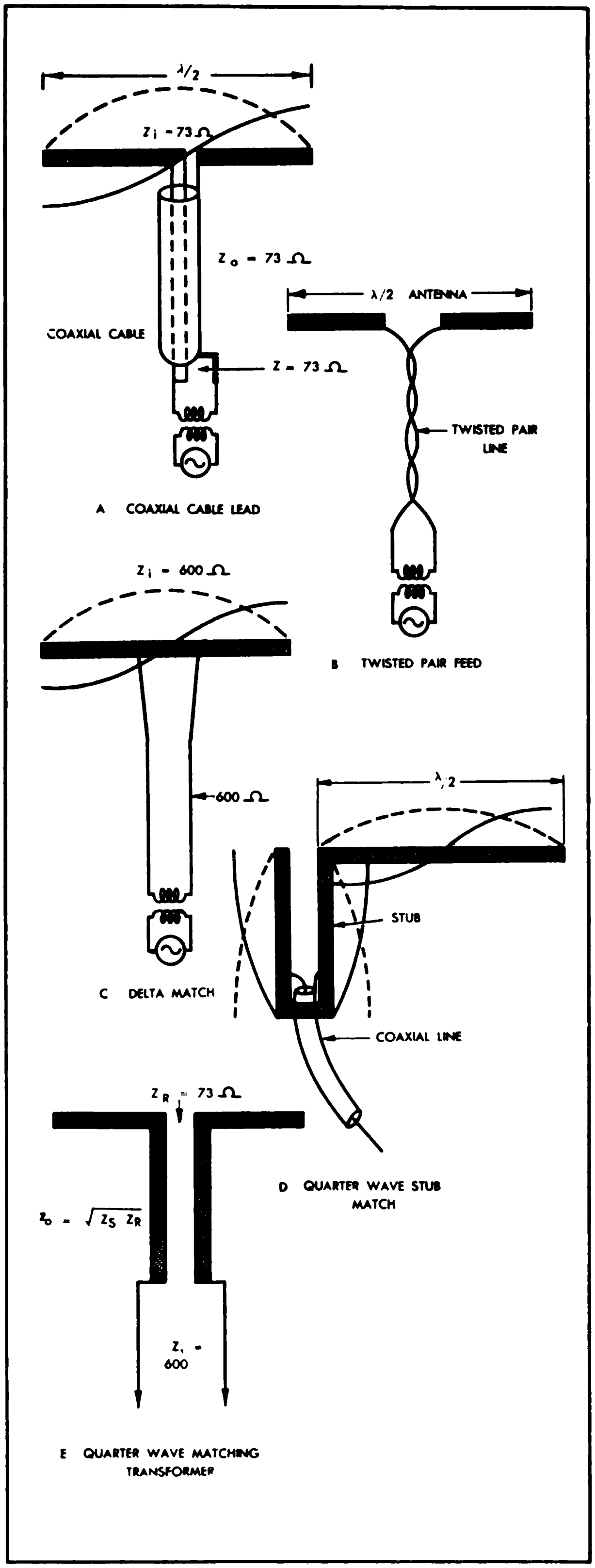
to antennas and to lines of any length — providing both are of resonant length. The line which is connected to the antenna can be a two-wire line, a coaxial line, or a waveguide. In radar equipment, the coaxial line or waveguide is preferred.

One advantage of the resonant line is that it makes matching devices unnecessary. In addition, it makes it possible to compensate for incorrect antenna lengths by transmitter tuning. Its disadvantages are increased I^2R losses due to high standing waves of current on the resonant feed lines, increased possibility of arc-over between lines due to standing waves on it, very critical lengths, and production of radiation fields. Radiation fields are undesirable because they distort the normal radiation pattern of the antenna.

Feeding Antennas with Nonresonant Lines

The illustration to the left on page 424 shows methods of exciting a half-wave antenna using nonresonant lines. Since the input to the center of the antenna at A is 73 ohms and since a coaxial cable with the highest Q has a characteristic impedance of 73 ohms, the most common method to feed this antenna is through a coaxial cable connected directly to the center of the antenna. This method of connection produces no standing waves on the line (coaxial cable) when a generator is matched to the line. Coupling to the generator is usually made, as you see at A, through a simple untuned transformer secondary. Another means of efficient transfer of energy to the antenna can be accomplished by a twisted pair line as shown at B. In a twisted pair, the impedance is about 70 ohms. It, like the coaxial cable, provides a good impedance match.

When a line does not match the impedance of the antenna, it is necessary to use special impedance matching devices. An example of a type of impedance matching device is the delta match shown at C. It gets its name from the fact that the wires form the Greek letter delta at the antenna. The delta match is used to match a line having more than 73 ohms impedance to the center of an antenna. The line in the example shown is a 600-ohm line. The connections from the line are spread apart at the antenna until the impedance between connections is equal to 600 ohms. In the delta match as the connections are moved away from the center, the voltage becomes higher and the current becomes lower. The ratio between these, known as the E/I ratio, increases from a very low value at the center to several thousand ohms at the ends. The two connections are set at the point where this ratio is equal to 600 ohms. This is the same as matching various impedances along a shorted quarter-wave section of line.



Feeding Antennas with Nonresonant Lines

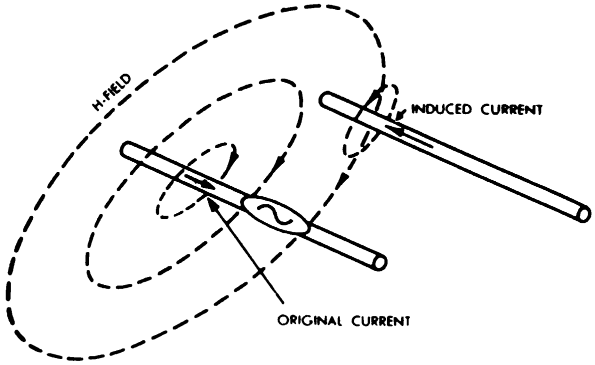
Another method of matching impedance is the quarter-wave stub match shown at D. In the quarter-wave stub match, the high impedance at the end of the antenna matches the open end of the stub. The impedance on the stub varies from zero at the short circuit to several thousand ohms at the open end. This makes it possible to connect a 70-ohm coaxial line a short distance from the shorted end at the 70-ohm point. It is possible to match almost any impedance along the length of the stub.

The quarter-wave transformer shown at E is another impedance matching device. This device is used for matching the low impedance at the antenna to the line of higher impedance. In the example shown, the characteristic impedance of the quarter-wave section is equal to 210 ohms. With this matching device, standing waves will exist on the antenna and the quarter-wave section but not on the 600-ohm line itself.

Nonresonant lines are characterized by small radiation, low voltage at all points for any given power, and noncritical length. In addition, they are preferred for the transfer of power over distances greater than one or two wavelengths.

Exciting Antennas by Radiation Fields

The antenna may be excited either by a direct connection or by radiation (induction fields) from nearby antennas. A half-wave antenna driven by a directly connected generator is shown in the illustration below. The flux from the driven antenna expands and cuts the other element and induces in it a current which varies at ~~a frequency nearly equal to~~ the



Exciting Antenna by Radiation Field

exciting frequency of the driven antennas and which, in turn, causes additional radiation fields. When the antenna elements are properly designed, the radiation fields resulting from the induced current will be nearly as strong as the original radiation fields.

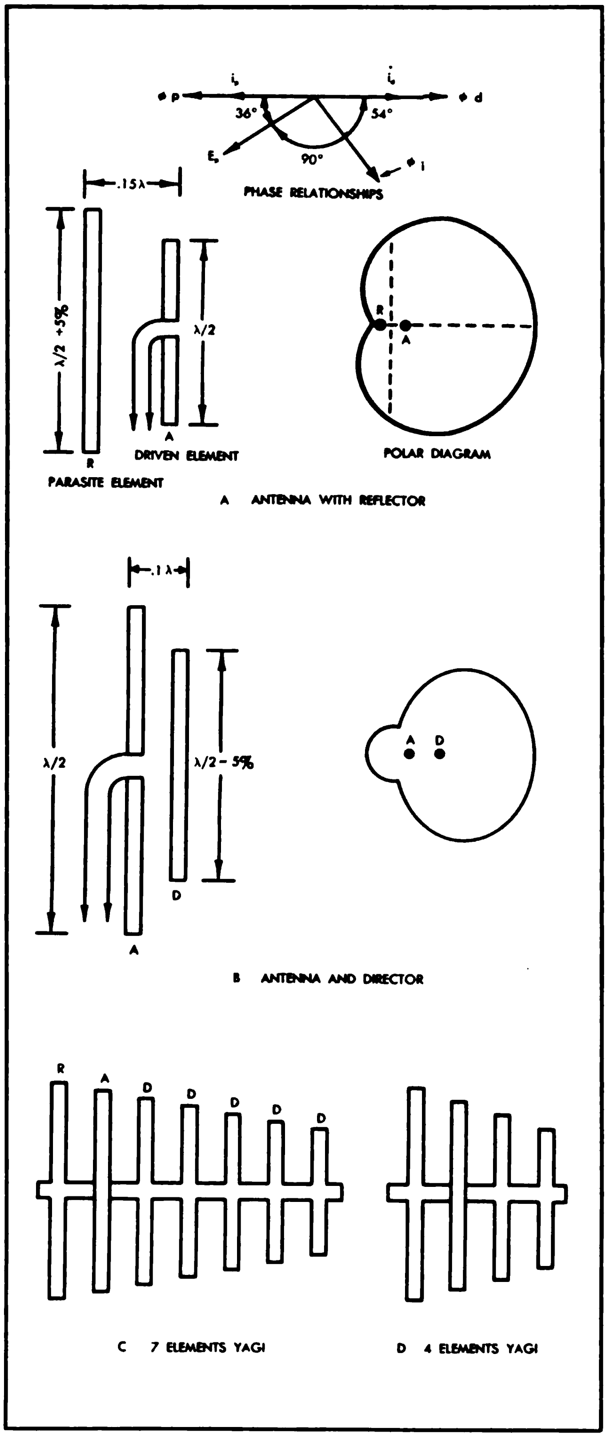
Besides being excited by a nearby antenna, an antenna may also be excited by radiation fields from a distant transmitting antenna. This takes place in any receiving antenna. In a receiving antenna, the current that is conducted to the receiver is the result of current induced by radiation fields generated by the transmitting antenna.

PARASITIC ANTENNA ARRAYS

Some uses for antennas require that the radiation be absolutely nondirectional while other uses require that the antenna be very sharply directional. An antenna system that is directional may consist of two or more simple half-wave elements so spaced that the fields from the elements add in some directions and cancel in others. The set of antenna elements is called an antenna array. There are two types of antenna arrays—driven arrays and parasitic arrays. In this manual we will discuss only the parasitic.

A parasitic array is an antenna system consisting of two or more elements in which only one of the elements is driven. The other element or elements are excited by induction and radiation fields produced by the driven element. With parasitic arrays, it is possible to obtain highly directional patterns.

The action in a parasitic array may be compared to the action of a transformer in which the primary current induces a current in the secondary, and the current in the secondary produces a magnetic field. This magnetic field, in turn, induces current back into the primary. Its action differs from that in a transformer in that the phase relationship between elements in an array varies according to the spacing between the elements. The elements are usually spaced an appreciable part of a wavelength apart. This makes it possible for the phase difference to vary over a range of more than 80°.



Parasitic Arrays

Two-Element Arrays

In a two-element parasitic array, as shown at A in the illustration above, the driven element is cut at the center for connecting a low impedance feedline. The length of the driven element is a half wavelength. This makes it self-resonant. The parasitic element, which is 5% longer than the driven element, is located 0.15 wavelength from it.

The polar diagram for two elements spaced 0.15 λ apart and excited out of phase is also shown at A on page 425. Note that most of the radiation is on the side of the driven element away from the parasitic element. Very little occurs on the side of the parasitic element, which acts somewhat as a reflector, since the radiation pattern is strongest in the direction of the driven element. Because of this action, the parasitic element is called a *reflector.*

The field passing from the reflector cuts the driven element and induces a voltage in it that changes the input current. The input impedance, which is a function of the current, is less than the 73 ohms for the antenna alone. It is usually about 50 ohms.

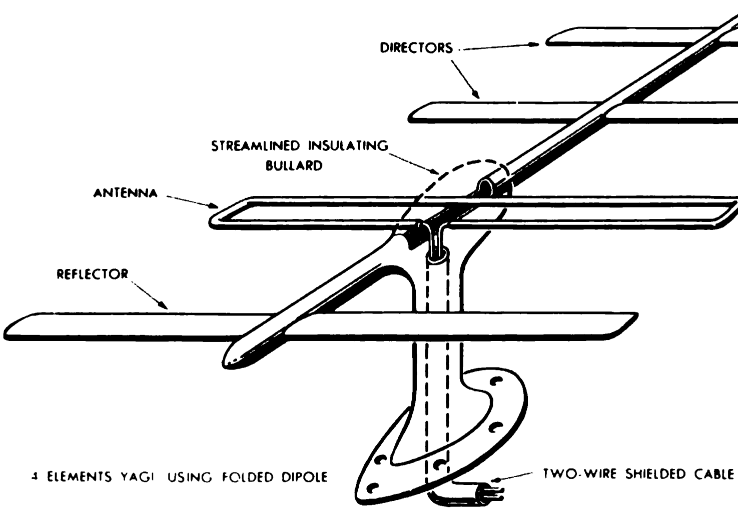
If the parasitic element is made shorter than the driven element and arranged as in B on page 425, it will act as a *director.* Most of the radiation will be in the direction of the parasitic element. The director is usually 5% shorter than the driven element and placed 0.1 λ from it. Sometimes the input impedance of this array is reduced to 20 ohms at the driven element.

In radar systems, the parasitic array usually consists of from 2 to 7 elements. Notice the 4-element Yagi array shown in the illustration below. In it, the antenna or driven element is insulated; reflector and directors are welded to a piece of tubing that runs parallel to the direction of propagation. This array has a beam-width of about 50°.

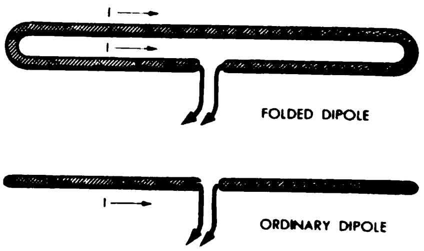
Directivity of Arrays

The directivity of an antenna array is expressed either in terms of the ratio of the power in the best lobe to the power radiated by a simple half-wave antenna, or as the ratio of the power in the best lobe to the power in the opposite direction. The direction of the best lobe is usually called the *forward* direction, while that in the opposite direction is called the *backward* direction. In other words, directivity is the ratio of the power in the forward field to that in the backward field. For example, the *front-to-back* ratio of the antenna array shown at B, in the illustration on page 425, is about 5:1. In decibels, this represents approximately a 7-db gain. The ratio of the forward power to the power from a single half-wave antenna is about 4 db. The gain of a two-element array using only a reflector is somewhat less than this gain. For this reason, two-element arrays used with radar equipment usually employ a driven element and a director.

Comparisons between the more directive types of arrays are made in terms of the beam angle.


DIRECTORS
STREAMLINED INSULATING BULLARD
ANTENNA
REFLECTOR
4 ELEMENTS YAGI USING FOLDED DIPOLE
TWO-WIRE SHIELDED CABLE

Yagi Array



Folded and Ordinary Dipoles

This angle is the angle between half power points in the main lobe. In arrays, the half power points are the points where the electric field strength is .707 as great as that along the axis of the beam.

Folded Dipole

In multi-element arrays, the input impedance drops as low as 15 ohms if an ordinary dipole is used. In these arrays, special matching devices are needed for matching this low impedance to the higher impedance of RF transmission lines. One method often employed in radar equipment uses a folded dipole as shown in the illustration at the bottom of page 426.

A folded dipole is a full-wavelength conductor which is folded to form a half-wave element. A better description is that it consists of a pair of half-wave elements connected together at the ends. In it the voltage at the ends of each element must be the same. In operation, the field from the driven element induces a current in the second element. This current is the same as the current in the driven element.

An ordinary dipole with a given current I produces a certain field intensity in space. Due to this field, there is also a certain power density per square meter in space. This power density is produced by the input power P. The relationship between the input resistance, the current, and the input power is expressed by the equation:

$$R = \frac{P}{I^2}$$

When the same current I exists in each of the two sections of the folded dipole, the field strength in space is doubled. This causes the power density per square meter to increase four times. In turn, the input power must be four times as great. In this case, to balance the equation, it is necessary to multiply R and P by 4 as follows:

$$\frac{4P}{I^2} = 4R$$

If each section of a dipole has the same diameter, the input resistance is four times that of the simple half-wave dipole. Increasing the diameter of one section makes the increase in impedance still greater. The input impedance to the driven element of a parasitic array drops to about a fourth of the value of the coaxial cable impedance, but the use of a folded dipole increases the impedance about four times. By using a folded dipole, a good impedance match is effected between the transmission line and the array.

RADAR ANTENNAS

A radar set is capable of detecting the presence of an object or target within the area which it scans, while at the same time measuring the range or distance of the target from the radar transmitter. Powerful signals sent out from the radar transmitter strike the target and are reflected back to the transmitter, where the echo or reflected signal is detected in a receiver. The time required for the signals to travel from the transmitter to the target and return to the receiver is used as a measure of the intervening distance.

A primary requirement of the radar system is that of illuminating the target with a maximum amount of signal so that a maximum reflected signal can be obtained. Thus, high-gain antenna arrays are needed. This reduces the power output required of the transmitter and makes for more efficient over-all design.

Another general requirement of the radar antenna is that the beam width of the antenna radiation pattern be as narrow as possible. This is a particular requirement of gunlaying radars used to position guns. The antenna should illuminate only a very small area at any one instant with a minimum amount of energy in the minor lobes. These minor lobes, if present, cause the reception of false echoes.

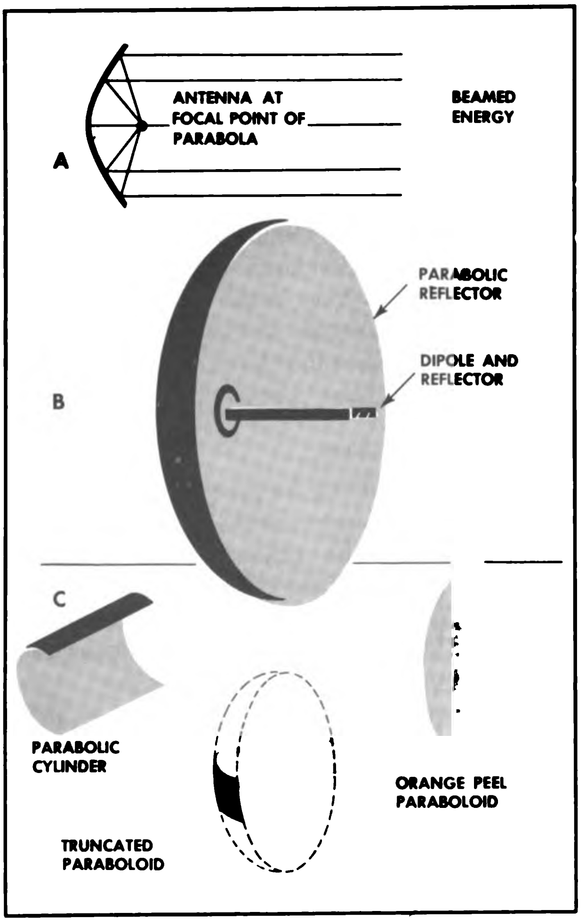
Parabolic Reflectors

As the transmitting frequencies approach the microwave region (where wavelength is measured in centimeters), the similarity between light waves and radio frequencies becomes more pronounced. Microwaves travel in straight lines and it is possible to reflect and to focus them in a manner similar to light waves.

One of the common methods used to direct light rays along narrow beams is through the use of parabolic reflectors. The reflectors, similar to the one in your flashlight, are shaped in such a manner that a source of light placed at the focal point is reflected as a narrow beam of light in a forward direction. A similar directing effect is employed in the transmission of microwaves.

A feed source, or antenna, is placed at the focal point of the parabolic reflector. The radiation from the source is reflected by the parabola in such a way that the energy leaves the antenna in a highly directional beam. Although light requires a polished mirrored surface for good reflection, microwaves are reflected by large metal surfaces. Thus, metal parabolas are used as microwave reflectors.

Part A of the illustration, Parabolic Reflectors, shows the formation of a narrow beam composed of parallel rays of energy when that energy originates at the focal point. All of the energy radiated from the antenna will be formed into a narrow beam and sent forward. But what of the energy radiated forward from the antenna itself? For example, suppose the feed to the parabolic reflector were a half-wave dipole. It would have a bidirectional pattern and, thus, would send out energy in two directions. One direction would be back into the parabolic reflector where it would be focused and sent forward in a narrow beam. The second direction would be directly forward from the half-wave antenna. This energy would have a wide beam width which would interfere with the desired narrow beam of the parabolic reflector.



Parabolic Reflectors

To avoid this second source of radiation, a simple half-wave parasitic reflector may be used in front of the driven antenna. Thus, by a combination of a parasitic element and a parabolic focusing element, a highly directional antenna array is achieved. One such arrangement is shown in part B of the illustration.

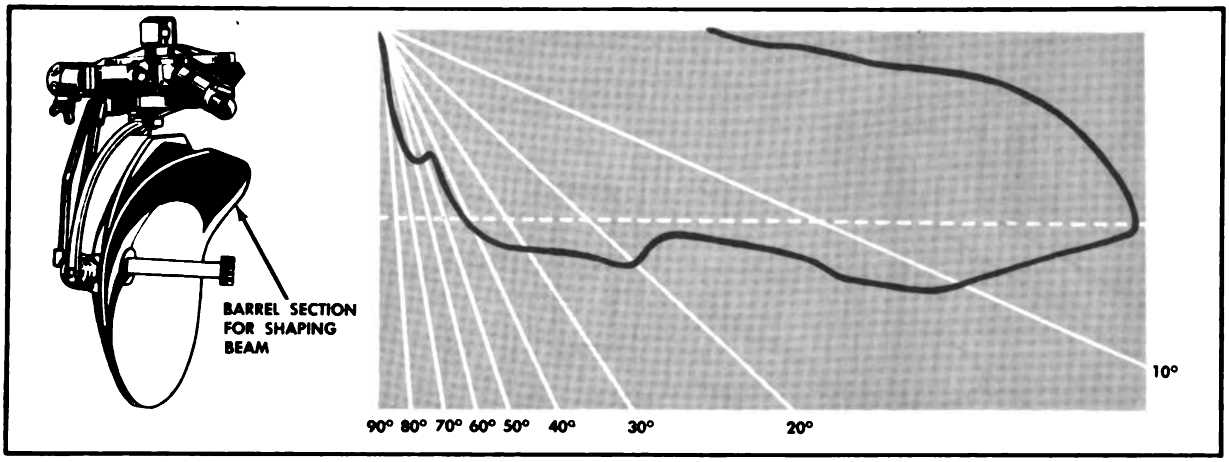
There are a variety of forms that the parabolic reflector may take. For example, consider the reflectors shown in part C. They illustrate possible variations in the type of parabolic reflectors that are in common use.

When you attempt to detect moving objects, it is advantageous to scan the desired area as rapidly as possible. One method of accomplishing this is to employ a fan-like beam which is narrow in the horizontal plane and wide in the vertical plane instead of a pencil-like beam. With the fan beam, a much greater area may be examined in a shorter time than is possible with a narrower beam. Also, this type of beam is useful for radar sets which are installed on moving objects. In this case, a sharp pencil-like beam might miss the object which is to be scanned.

A truncated paraboloidal reflector, as shown at C in the illustration to the left, may be employed to produce a vertical fan-like beam. For this reflector, the beam will be narrow in the horizontal plane, for the parabolic reflector will focus the energy properly in that plane. Since the reflector is truncated or cut so that it is shortened vertically, the resultant beam is allowed to spread in the vertical plane.

Cosecant-Squared Reflectors

In airborne radar sets that are designed for use in scanning a large ground area, a special beam shape is required. The beam should be narrow in the horizontal plane but broad in the vertical plane. In addition, the particular section on the ground that is being scanned must not be uniformly illuminated by the radar beam. Instead, the ground and objects on the ground under the aircraft should be illuminated by a small amount of power. This is required because echo signals from this location will be



Cosecant-Squared Antenna and Pattern

strong due to their closeness to the plane. On the other hand, a greater amount of power should fall on more distant objects in this sector because they normally produce the weakest echoes. Thus, the radiated field produced by the beam should increase in intensity at greater distances. In other words, the field intensity should vary as the cosecant of the angle between the horizontal and a line drawn between the aircraft and a given point on the ground. Since power varies as the square of the field intensity, then the required power pattern must be a cosecant squared pattern as shown in the illustration above.

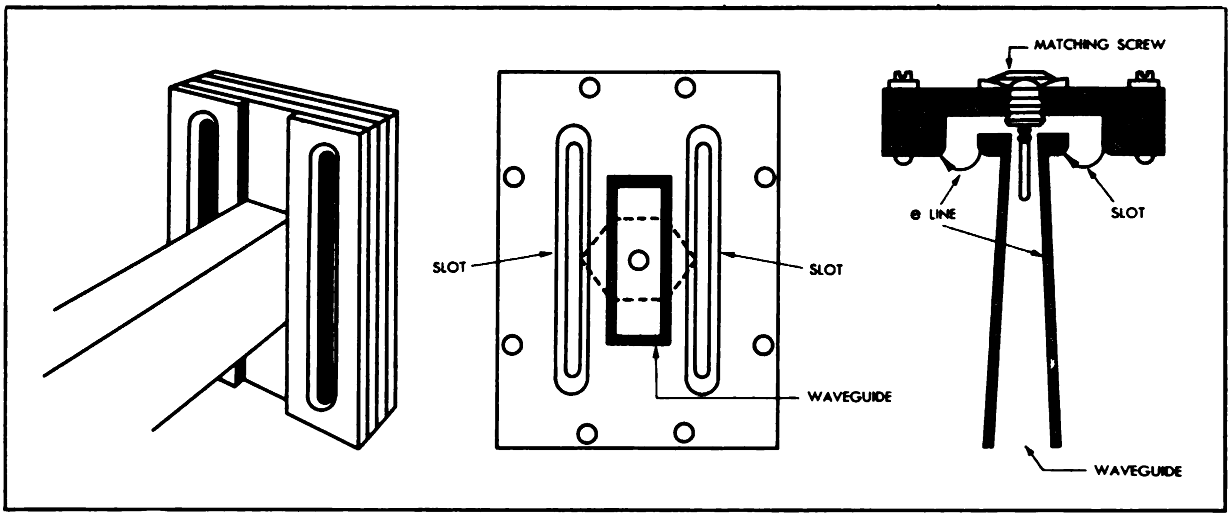
A common method of obtaining such a pattern is to use a special parabolic reflector in which the top portion is bent forward in order to produce the required pattern. Such a reflector, with the radiation pattern produced, is shown in the illustration above.

A TYPICAL 3-CM ANTENNA SYSTEM USING A PARABOLIC REFLECTOR

A typical example of a set with a 3-cm antenna system using a parabolic reflector is a radar navigation bombing set. As a navigational aid, this set is required to present a clearly defined picture of the terrain in all directions, both near and far, from the aircraft. This requires high definition in both azimuth and range. Good azimuth definition is provided by a sharply focused parabolic reflector. All directions are covered by rotating the antenna system through 360°. Good range definition is provided by a transmitted signal composed of a superhigh frequency carrier and a very short pulse. Presentation of near and far areas simultaneously is accomplished with a cosecant-squared reflector.

The entire antenna assembly is mounted on a rotating base which is turned by an electric motor. The system uses a 1 x ½ inch waveguide for an RF transmission line. A rotary waveguide joint, brings the signal through the rotating base. When the reflector is tilted vertically, the signal goes through a second rotary joint at the tilt axis. The waveguide is finally terminated in front of the reflector in a radiator, which in reality is a cavity at the end of the waveguide. The tapered section of waveguide in front of the reflector is called a Cutler feed.

As shown in the illustration on the next page, the radiator can be considered to be two small waveguides formed as a result of dividing the main waveguide into two halves and bending each half through 180 degrees. The fields emerge from openings or slots that are a half wave apart as shown below. The waveguide is tapered in the noncritical dimension to fit between the slots. The cavity impedance is matched to the waveguide by the matching screw. This adjustment is made at the factory and soldered in position. The radiated field is horizontally polarized. The design of the cavity is such that the energy striking the reflector decreases sinusoidally from the center toward each edge. Upon being reflected, the parabolic shape forms a horizontal radiation pattern which is only 3° wide at the half power points as shown in the illustration above.



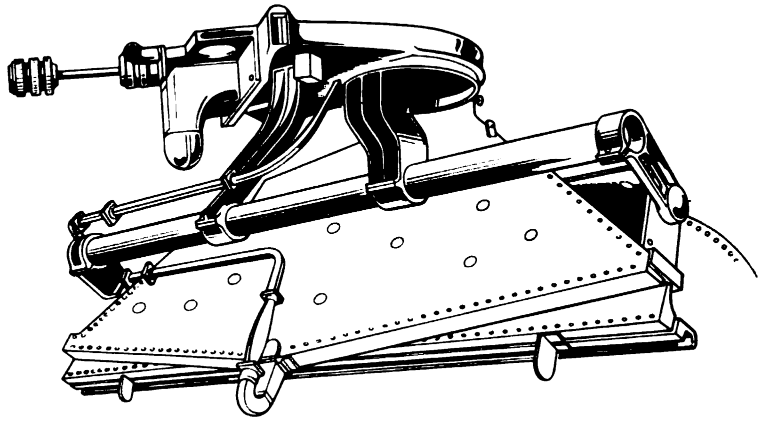
Cutler-feed Radiator

As mentioned before, it is necessary to obtain satisfactory indications from objects at different ranges with this set. On it there are a number of devices that make this possible. For example, when the adjustment is set for receiving nearby objects, the gain of the receiver is too low for the returns from distant objects to be visible. This is gotten around by using the specially shaped reflector which spreads the available RF field evenly over the entire usable range.

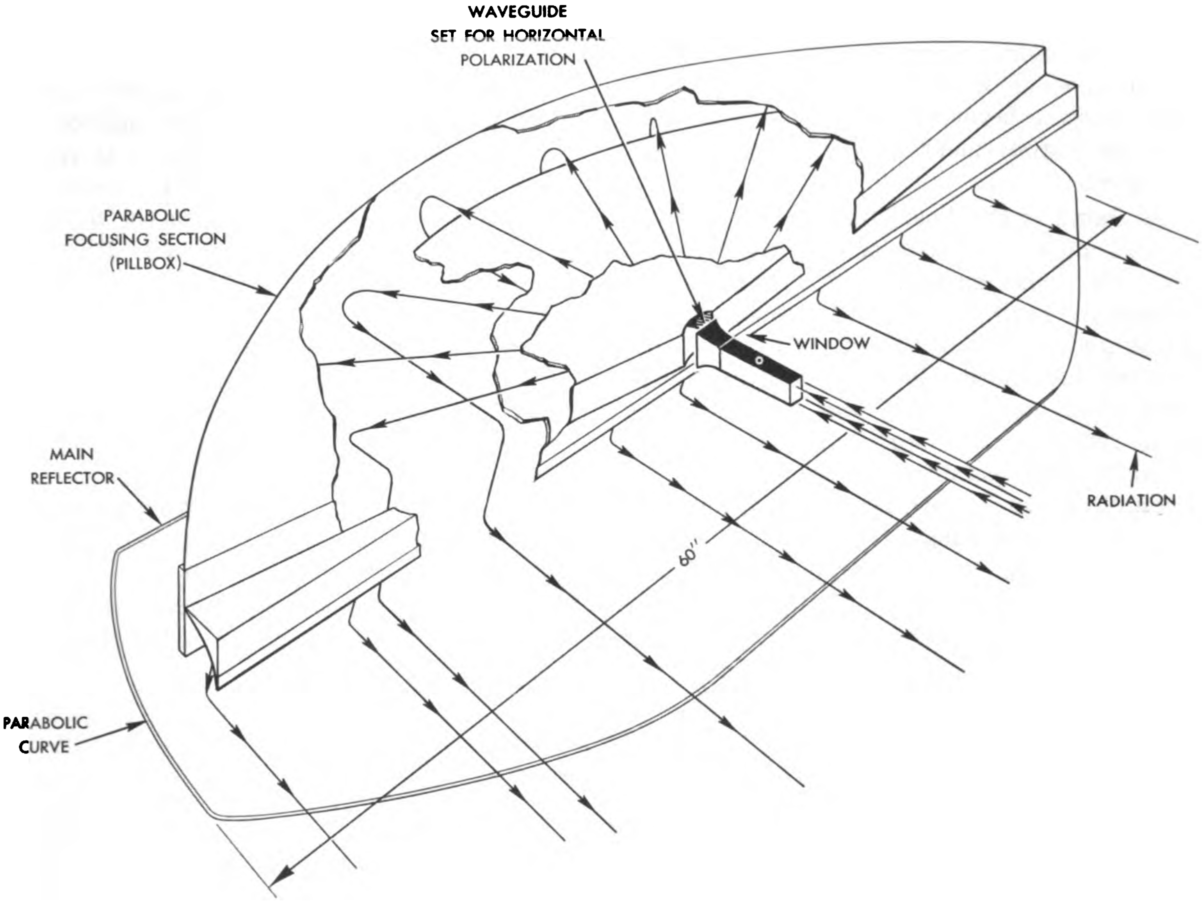
Usually, the antenna assembly is mounted below the fuselage on airplanes where the radiation is unobstructed in all directions. It is protected from the slipstream by a steamlined housing called the radome. The radome is constructed of material which causes low attenuation to fields which pass through it. Inferior radome materials will cause reflections from the inside surface. These radiations set up standing waves all the way back to the magnetron and cause it to change frequency. In addition, the reflections change as the antenna rotates.

THE AS-361 RADIATING SYSTEM

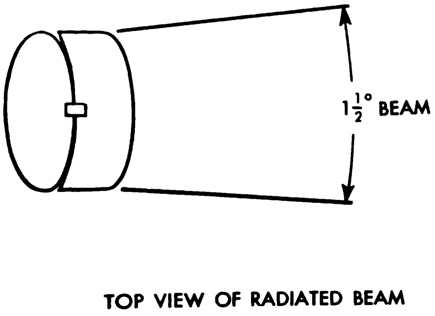
As previously discussed, a navigation and bombing radar system is required to present a clearly defined picture of the terrain in all

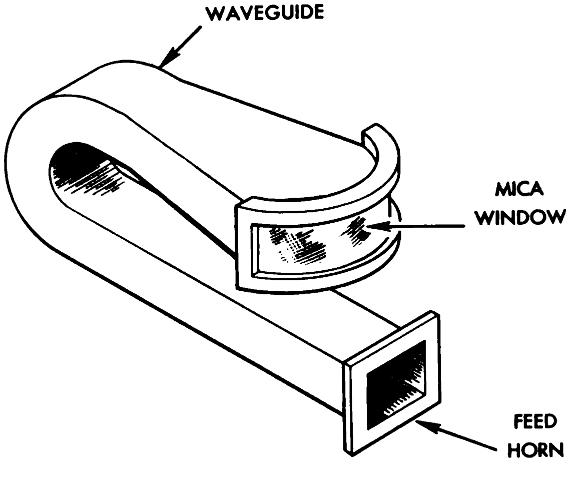


Antenna Assembly



Cutaway of Pillbox and Feed Horn of the AS-361





Radiating Assembly of the AS-361

directions. The narrower the beamwidth the better the definition in azimuth. The AS-361 antenna, which is used with several bombing systems has a beam width of 1.5° and is therefore an improvement over the system just discussed.

As shown in the illustration on page 430, the AS-361 antenna consists of two major assemblies, the mounting base, and a rotating table assembly. The drive unit and the rotating waveguide are mounted on the mounting base. The rotating table assembly consists of a rotating joint, feed horn, pillbox, and secondary and main reflectors. The whole assembly may be tilted by the operator in order to obtain an optimum picture regardless of changes in altitude or operating range.

The operation of this antenna can best be described by saying that the radiation pattern is a result of the combined action of three parabolic reflectors. In the illustration shown on page 431, you will notice that the energy is fed to the pillbox by a horn.

The horn is flared to match the impedance of the waveguide to that of the pillbox. This method of impedance matching was used in the cutler feed. The end of the horn is terminated with a mica window. This allows the waveguide to be pressurized and yet permits RF energy to pass through. A picture of the feed horn is also shown in the illustration on page 431.

The RF energy passes from the feed horn into the paraboloid pillbox. The wave fronts are then reflected in parallel paths to the secondary reflector, which redirects them to the main reflector. The main reflector refocuses the parallel wave fronts from the secondary reflector and reflects them out toward the target. The combined action of the pillbox, secondary reflector, and main reflector produces a cosecant squared pattern, 1.5° in azimuth and 55° in the vertical plane.

Chapter 38

A Radar System

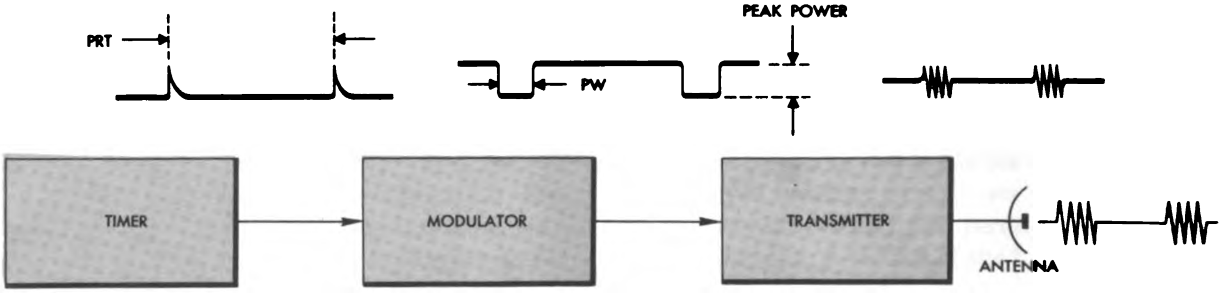
A radar system has three equally important channels. They are the transmission, reception and indicator channels. The *transmission* channel generates and radiates short pulses of RF energy. The *reception* channel collects the reflected pulses and prepares them for evaluation. The *indicator* channel generates a time base and displays the pulses on the time base. Having studied all the necessary circuits, you are now ready to combine them to form a basic pulse-modulated transmission channel.

RADAR TRANSMISSION CHANNEL

The illustration at the top of the next page is a simplified block diagram of a pulse-modulated transmission channel. The function of the timer is to trigger the transmitter (and the sweep) at the proper time intervals (PRT). It therefore determines the pulse recurrence frequency (PRF). The modulator receives these trigger pulses and converts them to high-powered rectangular pulses of a definite time duration (PW) as shown in the illustration. These rectangular pulses trigger the transmitter which produces the high frequency pulses. These pulses are then radiated by the antenna.

The simplified block diagram may be expanded into a more complete block diagram of the transmitter channel of the AN/APS-42 or the AN/APS-23 as shown in the illustration below it. The timer section shown in the diagram is the one used in the APS-42, consisting of a single blocking oscillator. The APS-23 timer uses two additional stages, but in both cases, the function of the timer is the same — to determine the PRF of the set.

The pips from the timer trigger the thyratron, allowing the pulse line to discharge, giving the desired pulse width. The pulse line discharges through the primary winding of the pulse transformer, through the thyratron and back to the other side of the pulse line. The changing current in the primary of the pulse transformer induces a 12,000-volt, negative-going square wave across



Simplified Block Diagram of Pulse-modulated Transmission Channel

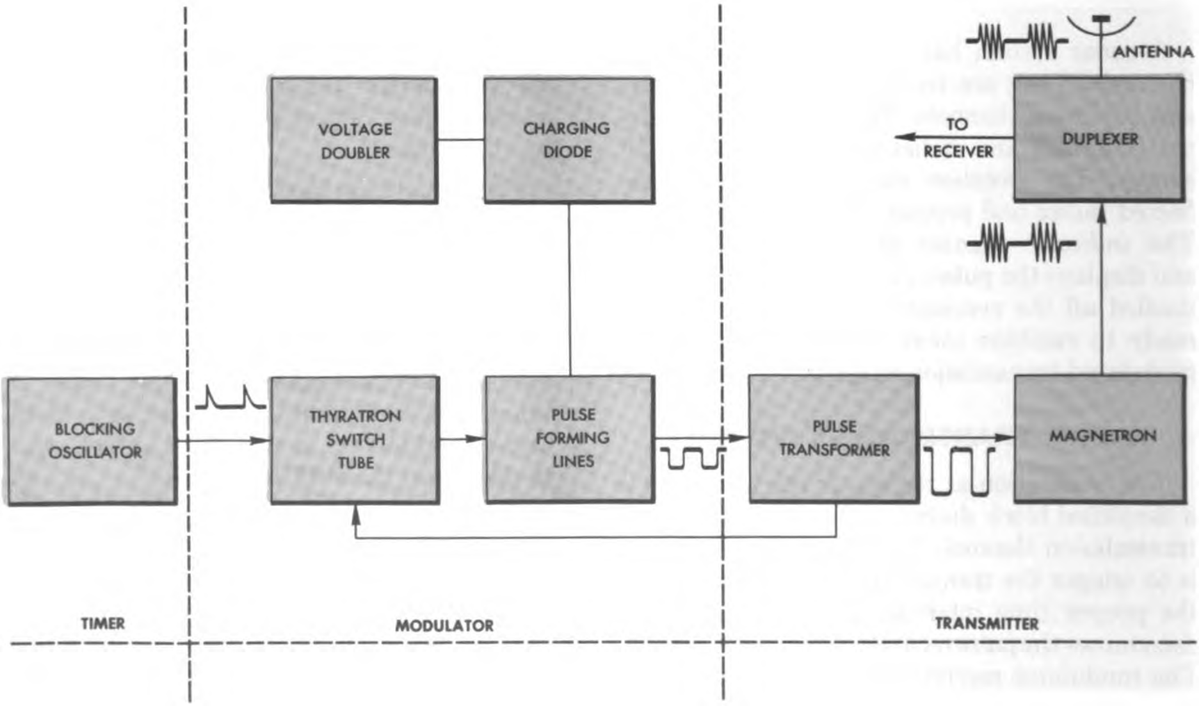
the secondary of the transformer. This high-voltage pulse shocks the magnetron into operation and it oscillates at approximately 9375 megacycles for the duration of the high-voltage pulse.

AN/APS-42 TRANSMISSION CHANNEL

The transmission channel of the AN/APS-42 may be analyzed with the aid of the schematic diagram on page 436 The three sections (the timer, the modulator, and the transmitter) are discussed in the following paragraphs.

Timer Section

The master timer produces trigger pulses at a predetermined pulse repetition frequency. It is a free-running blocking oscillator and produces very narrow pulses. For the theory of operation of a blocking oscillator, refer to chapter 31. The output frequency of the blocking oscillator (V_1) is selected by the PRF switch. This switch selects the proper RC circuit to produce the desired PRF. The APS-42 has three pulse repetition frequencies available (800 cps, 300 cps, and 200 cps), and the proper



Block Diagram of Transmitter Channel

PRF is switched in automatically according to the radar operator's control settings. The resistors R_1, R_2, and R_3 are variable so that very accurate adjustments of the PRF can be made.

Modulator Section

The modulator consists of pulse-forming network Z_1, thyratron switch tube V_2, charging diode V_3, clipper diode V_4, charging choke L_1, and the modulator high voltage supply (a voltage doubler whose output is 2500v DC).

When power is applied to the set, current to charge the pulse-forming network flows from ground through the primary of the pulse transformer, through the pulse network, through the charging diode V_3, the charging choke L_1 through the power supply back to ground. The pulse network is then charged to 2500 volts with the polarity shown. When the pulse network becomes fully charged, the magnetic field built up in the charging choke collapses. By this self-inductance, a voltage approximately twice the power supply voltage is established and the pulse-forming network is held at this high potential. The thyratron V_2 and the clipper diode V_4 prevent the discharge of the capacitors within the pulse-forming network. Since these tubes are cut off between pulses from the timer, they appear as infinite impedances and do not provide a discharge path for the network.

When the master timer V_1 conducts, it supplies a trigger voltage to the grid of the thyratron. The switch tube ionizes and provides a closed circuit through which the pulse-forming network discharges. This discharge continues for a period of time dependent upon which pulse line is in the circuit. In the illustration shown, the ¾μs and the 1½μs pulse lines are switched in by relay K_1, so that a 2¼μs pulse width is established.

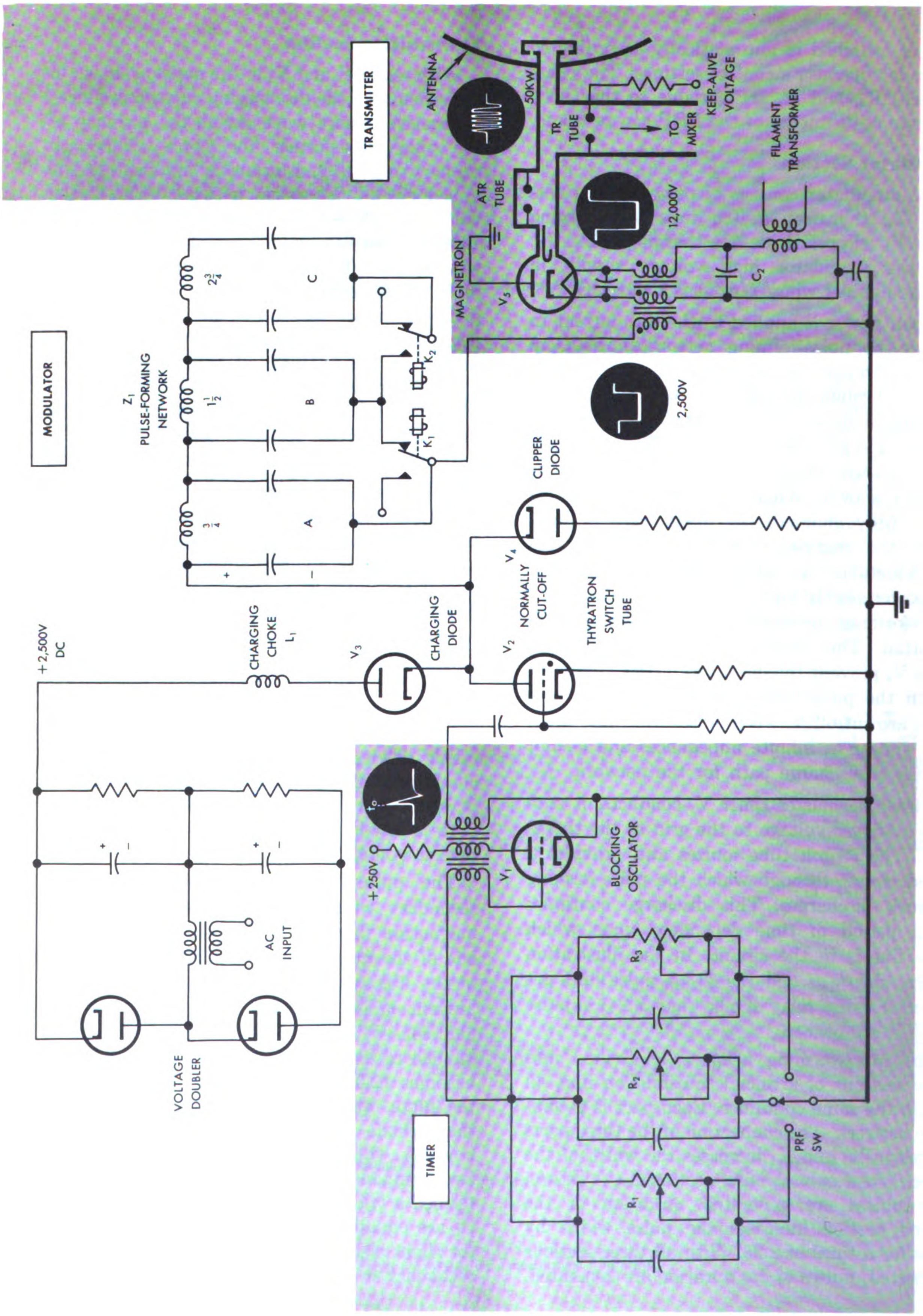
The pulse-forming network is divided into three sections of artificial transmission line, all within the same container. Each part is a complete network. The connection of the three parts by relays K_1 and K_2 increases the width of the pulse. When only the first section Z_{1A} is in use, ¾μs pulses are generated. Addition of the second section increases the pulse width by 1.5μs for a total of 2.25. With all three sections connected, pulses of 5μs duration are generated.

Transmitter Section

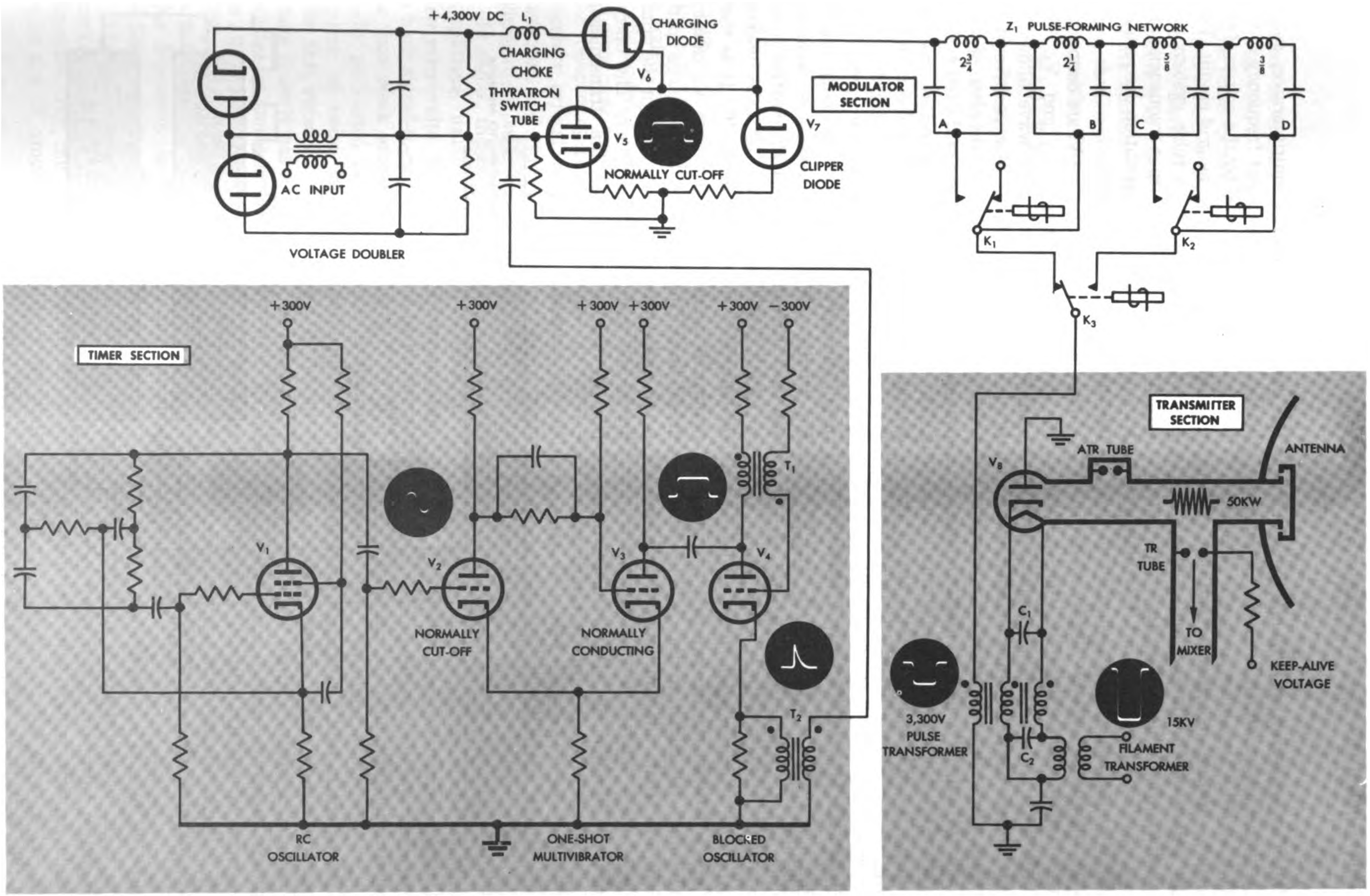
With relays K_1 and K_2 in the position shown in the schematic, the discharging current produces a negative square wave across the primary of the pulse transformer for 2¼μS. Since the primary of the transformer is matched to the impedance of the pulse network (50Ω), the voltage divides equally between these two units. Thus the square wave impressed on the primary winding is 2500 volts in amplitude. The turns ratio of the pulse transformer is five to one, giving a voltage step-up of one to five. Thus a negative pulse of approximately 12 kilovolts is placed on the cathode of the magnetron V_5. The peak power of this pulse is on the order of 150 kilowatts, but since the efficiency of the magnetron is only 30 to 40 percent, the power output of the magnetron is about 50 kilowatts.

Arcing occasionally occurs within the magnetron. This condition causes momentary pulse-transformer impedance mismatching, and a small voltage of polarity opposite to that of the modulator pulse may appear across the pulse-forming network after the pulse. This small reverse voltage is in series with the power supply voltage at the start of the next charging cycle. Hence the total voltage available to produce the charging cycle is increased, and the pulse-line voltage may build up to a value great enough to break down the thyratron. To prevent this condition, clipper diode V_4 is connected in parallel with the thyratron. This diode provides a low-impedance path to ground for the reverse voltage.

The pulse transformer has two identical secondary windings, one in each side of the magnetron filament voltage supply. The cathode of the magnetron is connected to one side of the filament line. The reason for this is to prevent arcing between the cathode and the filament. When the 12-kilovolt trigger pulse is applied to the secondary windings, the cathode becomes 12,000 volts negative with respect to the plate, and the magnetron operates. Since the cathode is connected to the filament line, however, there is no potential between these two components. This type of transformer is termed a *bifilar* transformer. Capacitors C_1 and C_2 are equalizing capacitors which compensate for slight electrical differences between the two secondary windings of the pulse transformer.



Schematic Diagram of APS-42 Transmission Channel



Schematic Diagram of APS-23 Transmission Channel

When the magnetron fires, the high-power RF energy ionizes both the ATR and TR tubes. Their reflected impedances are such that the waveguide appears continuous and the RF energy passes on to the antenna, where it is then radiated into space. When the transmitter pulse ceases, both the ATR and TR tubes deionize. Because the reflected signals do not develop sufficient potential to fire the tubes, the received energy is directed through the TR tube to the receiver.

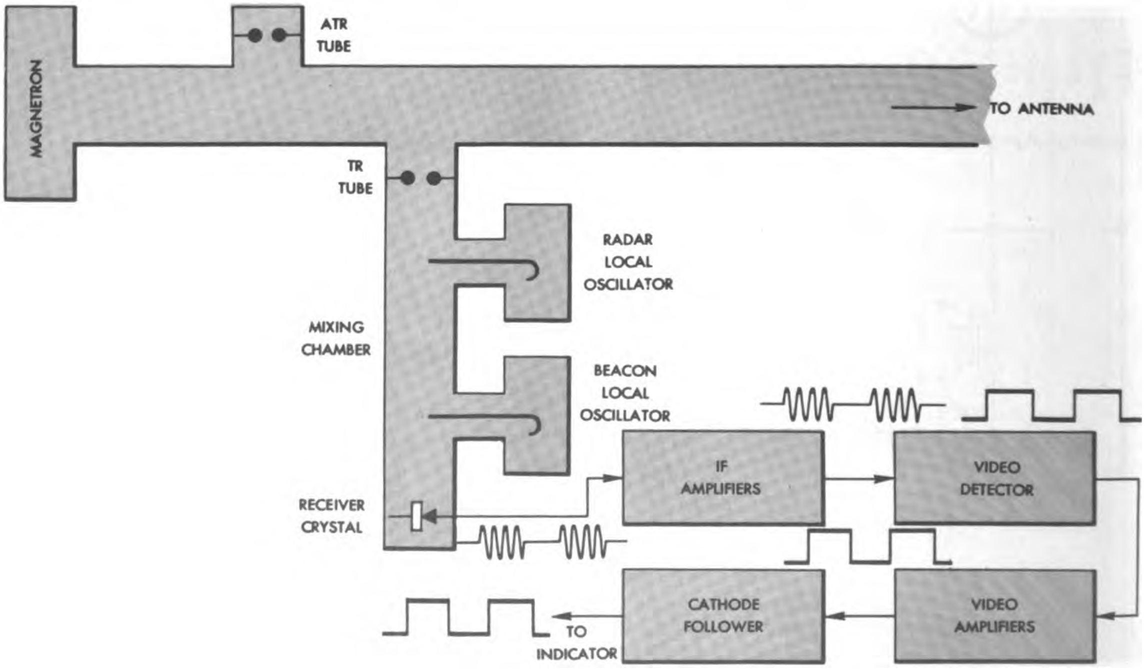
The TR tube has a 250-volt *keep-alive* voltage impressed across it at all times. This potential keeps the TR tube very near its ionization point. It therefore fires very readily during transmission and keeps the transmitted energy out of the receiver.

AN/APS-23 TRANSMISSION CHANNEL

The illustration on page 437 is a schematic diagram of the transmission channel of the AN/APS-23. Note that it is very similar to the AN/APS-42 transmission channel. The only differences between the two are that two additional stages are added to the APS-23 timer section and that different pulse-widths are available.

The timer section of the APS-23 consists of an RC oscillator, a one-shot multivibrator, and a blocked oscillator. The output frequency of the RC oscillator is determined by the values of the resistors and capacitors in the grid circuit. This value can be changed by a relay system (not shown) to produce three output frequencies corresponding to the pulse repetition frequencies desired (800, 300, or 200 cps).

The sine waves produced by V_1 are coupled to the one-shot multivibrator V_2 and V_3. The positive portion of each sine wave allows V_2 to conduct, cutting off V_3. The positive-going rectangular wave produced at the plate of V_3 is coupled through T_1 to the grid of the blocked oscillator. The transformer is wound so that the leading edge of this rectangular pulse is felt on the grid of V_4 as a negative potential, and does not trigger the stage. When the rectangular wave ends, however, the decrease in potential is felt on the grid of V_4 as a positive potential and fires the tube. A very sharp positive pulse is taken from the cathode of this blocked oscillator and stepped up across T_2. It is then applied to the grid of the thyratron V_5, causing it to fire. This produces a discharge path for the pulse network Z_1 down through the primary of the magnetron transformer. The operation of the remainder of the transmission



Basic Radar Reception Channel

channel in exactly the same as that of the APS-42.

The pulse-forming network of the APS-23 is divided into four sections of artificial transmission line. Each part is a complete network. The connection of the four parts by relays K_1, K_2, and K_3 increases the width of the pulse. We have, then, a choice of any one of four different pulse widths: ⅜µs, 1µs, 2¼µs, or 5µs.

RADAR RECEPTION CHANNEL

The function of the reception channel is to receive and detect target returns. As you have seen, the transmission channel generates and radiates bursts of radio frequency energy which travels out into space and hits distant targets. A portion of this RF energy is reflected to the receiving antenna. The receiver mixes this RF energy with RF energy from a beat frequency oscillator to produce a new lower frequency (the heterodyning principle). The receiver then amplifies this new frequency (intermediate frequency), removes the target information (intelligence), and applies this intelligence on to the indicator in the form of voltage pulses at a video frequency.

Basically, a radar reception channel consists of two local oscillators (one for radar and one for beacon operation), a mixing chamber, a receiver crystal, several IF amplifiers, a video detector, and video amplifiers, as shown in the illustration on the preceding page.

In radar, the received signal is of very small amplitude, usually in the neighborhood of a few microvolts. Because of this low amplitude, a high signal-to-noise ratio must be maintained so that weak echoes are not obscured by internally generated noise. The mixer is the main source of noise, since it is the input stage. For this reason a crystal is used as the mixer stage since it has a very low noise output. The crystal is non-selective and will pass current at all four of the above-mentioned frequencies. Still, the small input circuit to the IF amplifiers is tuned to 60 megacycles; consequently, only this frequency is passed on to the remainder of the receiver.

These frequencies pass on to the receiver crystal. Since a crystal passes current in one direction only, the pulses are rectified and passed on to the IF amplifiers as positive-going bursts of RF energy. The coupling circuits of the IF amplifiers are tuned to 60 megacycles; therefore, only the intermediate frequency is amplified to any extent.

The IF signal, after passing through several IF amplifiers, is applied to the video detector. In this stage, the radio frequency energy is converted into energy at a video frequency. The resulting waveshape is a series of positive-going voltage pulses. The pulses, each of which represents a target return, are amplified through a series of video amplifiers and applied to the indicator where they are displayed as visual target returns.

AN/APS-23 RECEIVER

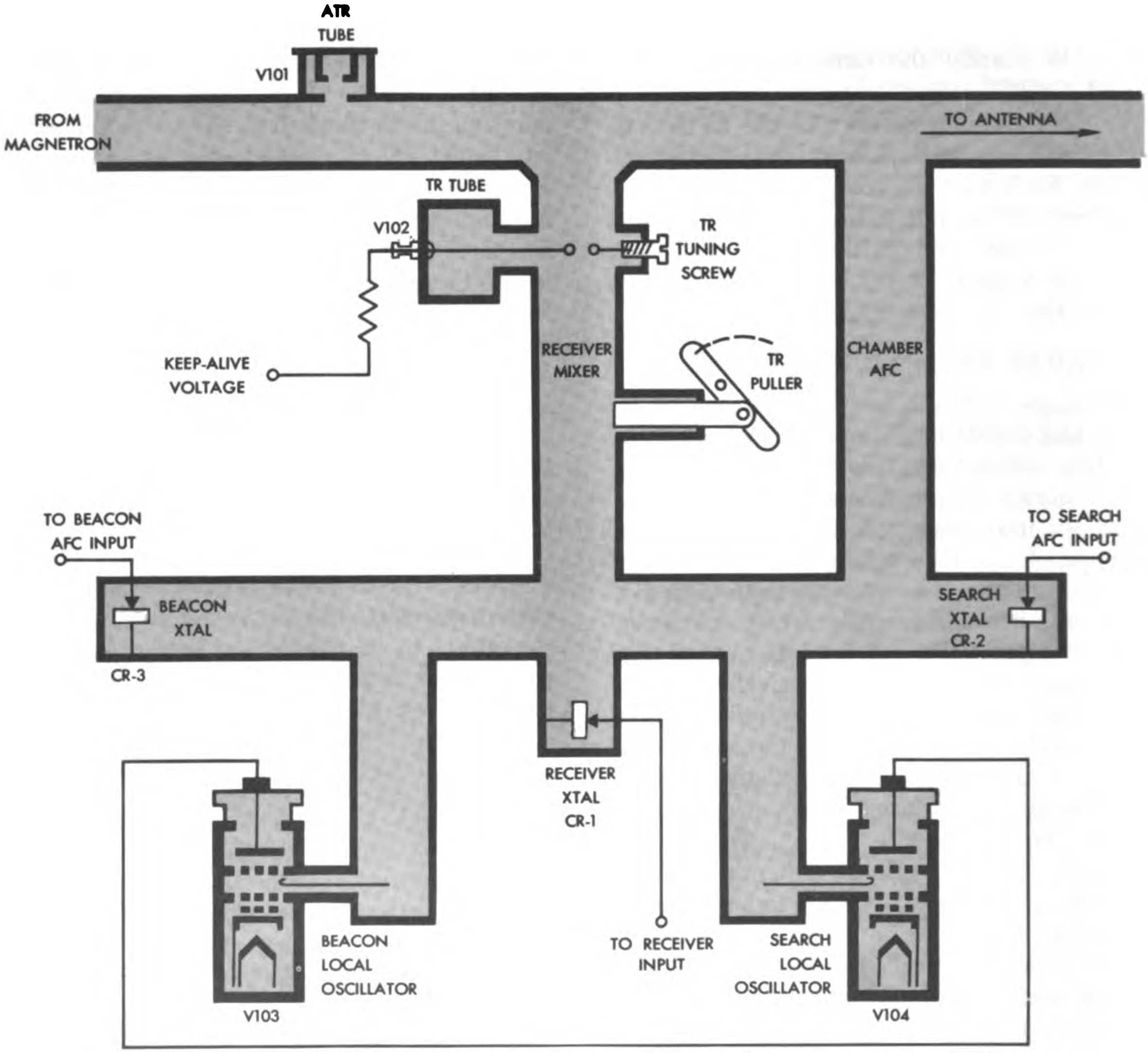
Although the reception channel described is a practical circuit and is still used in some of the older radar sets, the target presentation leaves something to be desired. In later sets, certain refinements have been incorporated in this basic receiver to achieve more effective target presentation. The AN/APS-23 is an example of a radar set whose receiver incorporates some special features.

The Mixer

The transmitter sends out radar pulses at a regular rate. In the time between pulses, radar echoes return from targets. These echo signals are picked up by the antenna and directed by the duplexer into the mixer. The mixer consists of a rectangular block containing four interconnected chambers as shown in the illustration on the next page. The four chambers are: the search local oscillator chamber, the beacon local oscillator chamber, the receiver-mixer chamber, and the automatic frequency control (AFC) chamber.

SEARCH OPERATION. Part of the search local oscillator output is fed to the receiver mixing chamber. Here (in the crystal) it is mixed with the echo signal (the radar return) coming in through the TR tube. Of the frequencies produced, only the 60 megacycle signal is passed to the remainder of the receiver by the IF preamplifiers.

During transmission, a part of the local oscillator output is coupled to the AFC mixing chamber where it is mixed with a portion of the transmitter energy coming through choke Z_1. The mixed output is fed through the Search



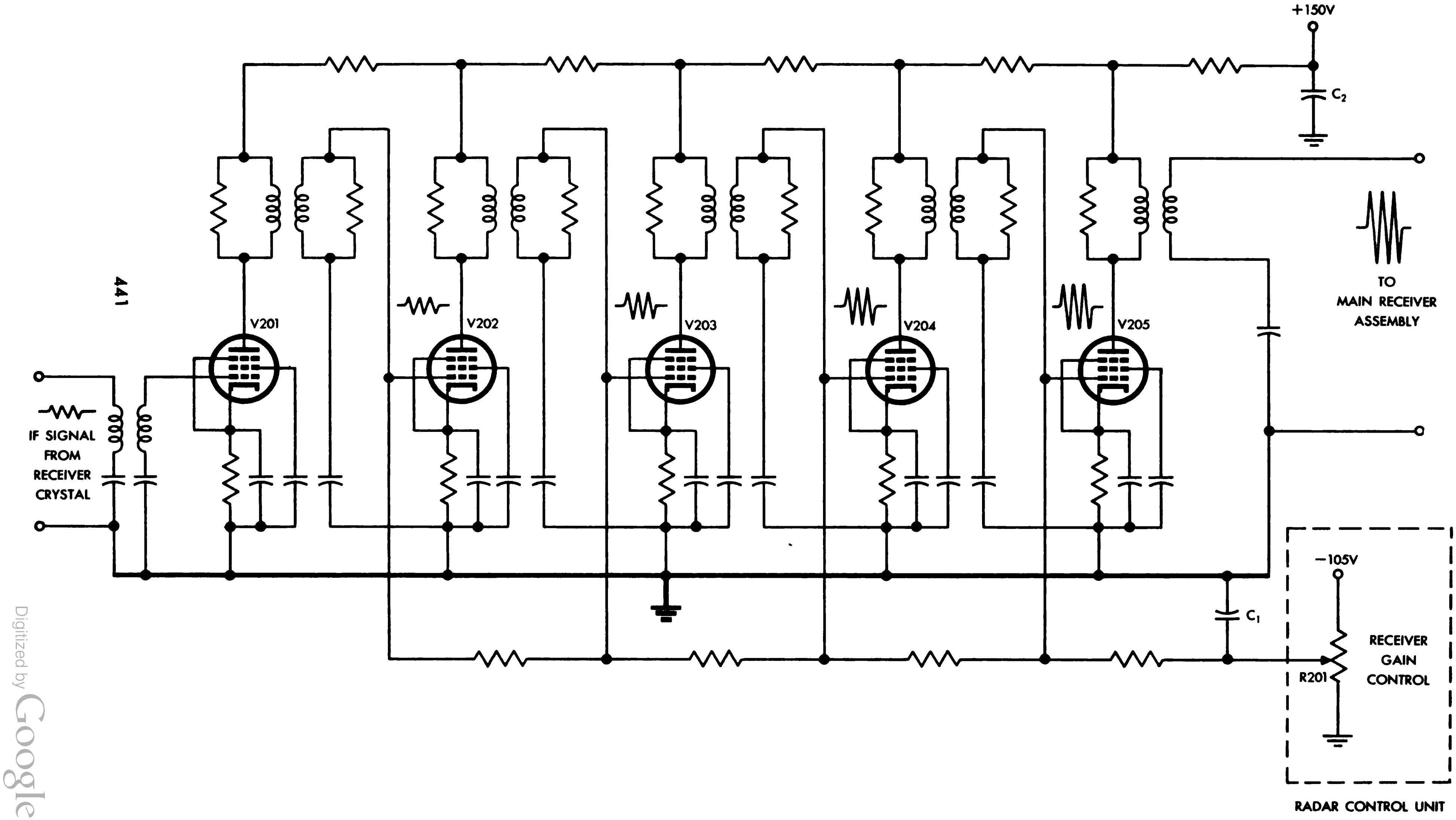
APS-23 Mixer and Local Oscillators

AFC crystal CR-2 to the AFC circuit. The AFC circuit holds the frequency of the search local oscillator 60 megacycles below that of the transmitter.

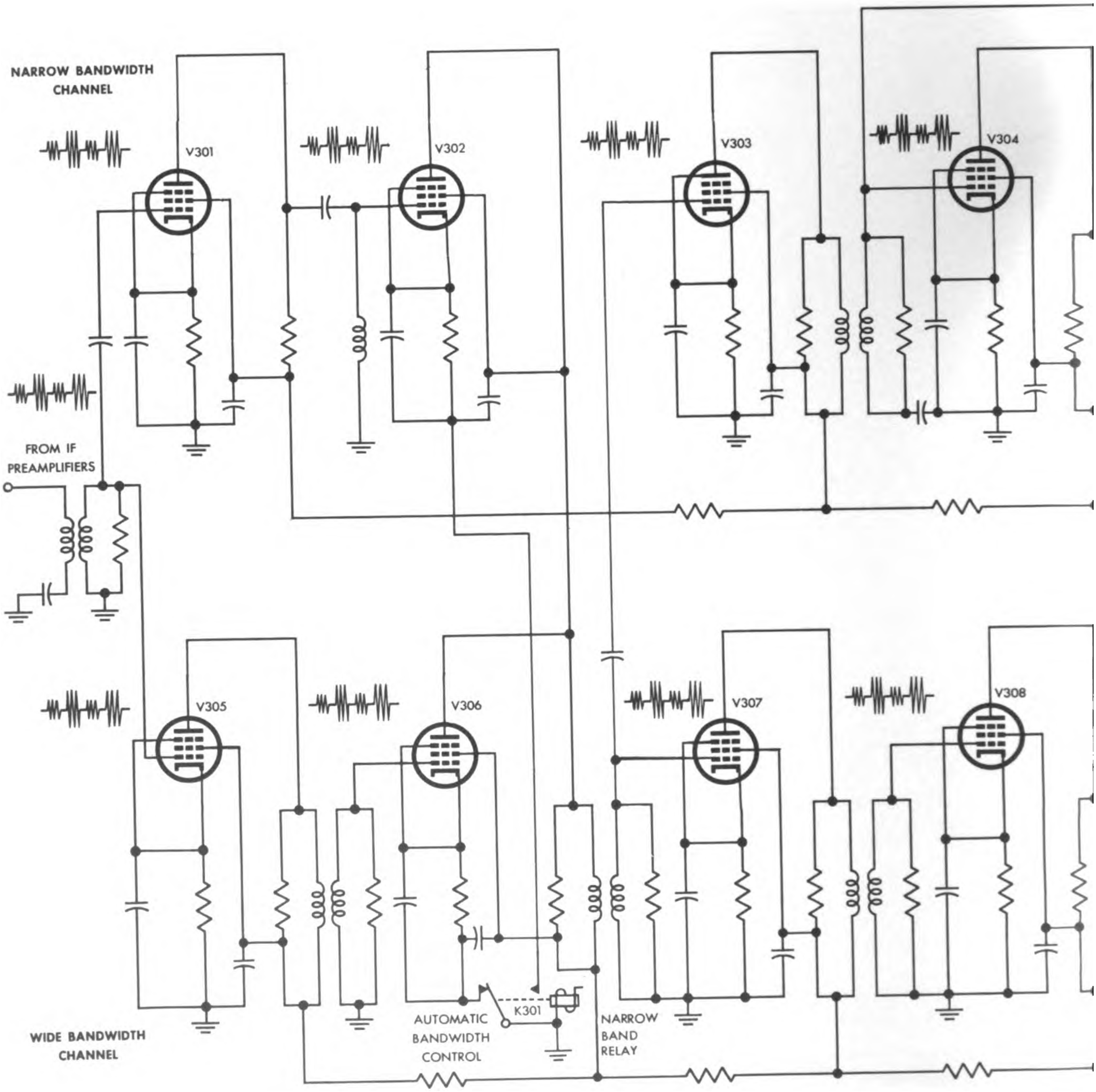
BEACON OPERATION. Part of the beacon local oscillator output is coupled to the receiver mixer where it is mixed with the beacon signal coming through the TR tube. The difference frequency is then fed through CR-1 and on to the receiver IF amplifiers. The beacon local oscillator is held by the Beacon AFC circuit at the beacon cavity frequency. The beacon cavity is tuned to 9250 megacycles, which is 60 megacycles below the beacon transmitter frequency. The TR puller tunes the TR box to 9310 megacycles to permit the beacon signals to pass through.

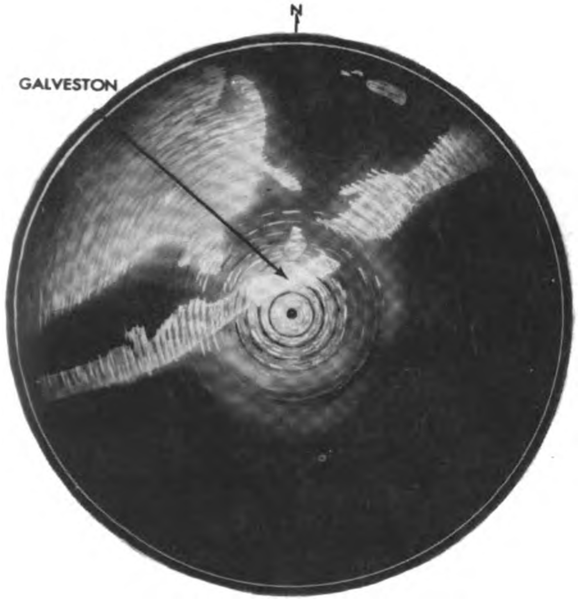
The Local Oscillator

The search (or beacon) local oscillator generates an RF output which mixes with the echo (or beacon) signals to give the correct 60-megacycle IF. Two separate oscillators are used, one for beacon operation and one for search operation. The local oscillator output frequency is controlled by the repeller plate voltage which can be varied between −100 and −200 volts. In AFC operation, the repeller plate is connected to the output of the AFC circuits which maintain the proper local oscillator frequency. If, for any reason, you do not desire to use AFC, either local oscillator may be tuned manually by the MANUAL TUNING



IF Preamplifier Assembly — Simplified Schematic


NARROW BANDWIDTH CHANNEL
V301
V302
V303
V304
FROM IF PREAMPLIFIERS
V305
V306
V307
V308
AUTOMATIC BANDWIDTH CONTROL
K301
NARROW BAND RELAY
WIDE BANDWIDTH CHANNEL



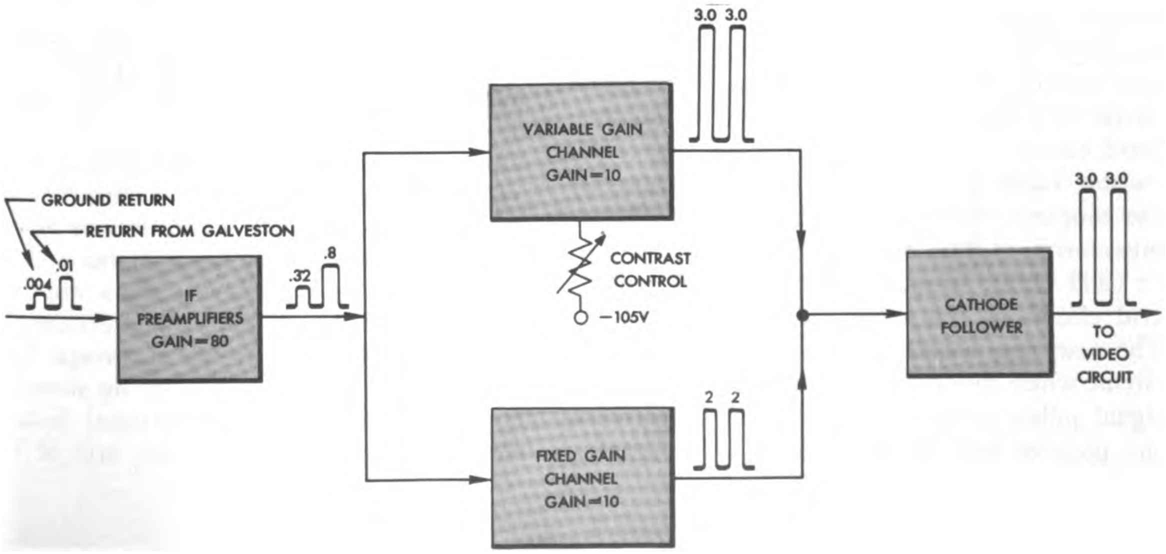
Approaching Galveston

which limits the output from Channel One to three volts. The last stage in Channel Two, also a video limiter, limits the output of Channel Two to two volts. These two outputs are combined in V312, a cathode follower, and the output from V312 is 3V for each signal.

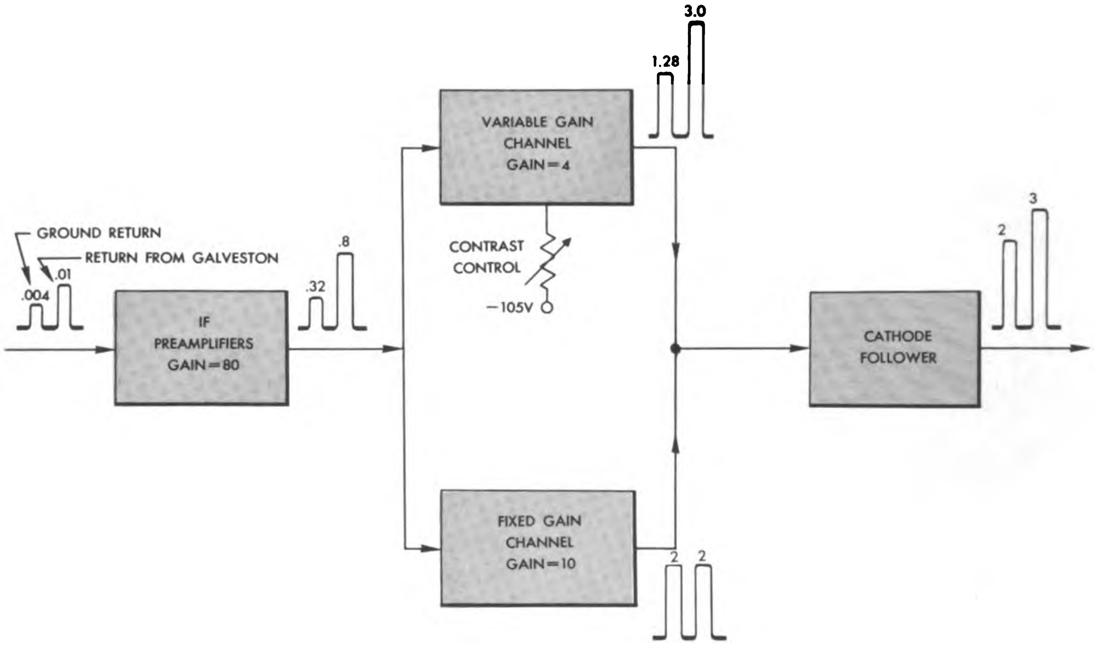
See the schematic on Page 445. The signal on the left-hand grid of V312 is two volts and the signal on the right-hand grid is three volts. The cathode of V312 follows the grid (in this case the right-hand grid) and the signal on the cathode is three volts for both the mapping background and the Galveston return. This is true because the "three-volt grid" causes the cathode to become positive enough that the "two-volt grid" side is cut off. Since both the land return and the Galveston return are three volts, you will be unable to see Galveston against the land background. By setting in a given amount of Contrast, Galveston will come in and you will maintain the land-water contrast. The reason for this is shown in the diagram labelled "Operating With Contrast". The gain of the variable gain channel has been reduced to four by setting in CONTRAST. The return from Galveston is limited to three volts. In the fixed gain channel, both returns are amplified and then limited to two volts. When the signals combine in the cathode follower, the ground return becomes two volts and the return from Galveston three volts. The land-water contrast has been maintained and you can easily distinguish Galveston from the background returns, because Galveston will appear as a much brighter target return."

The use of CONTRAST is not restricted to land-water areas. Target-separation trouble will also be encountered over areas containing many extremely strong targets clustered together, as in a large city. By the proper settings of RCVR GAIN and CONTRAST, you can keep the target return intensity down to the desired level and at the same time retain the mapping background for reference.

AUTOMATIC PULSE STRETCHING. Pulse stretching is used to increase the length of the received beacon signals to give a clear presentation on the indicator screen. Pulse stretching is accomplished by changing the time constant of the video detector V309A and V309B. You can see



Operating Without Contrast



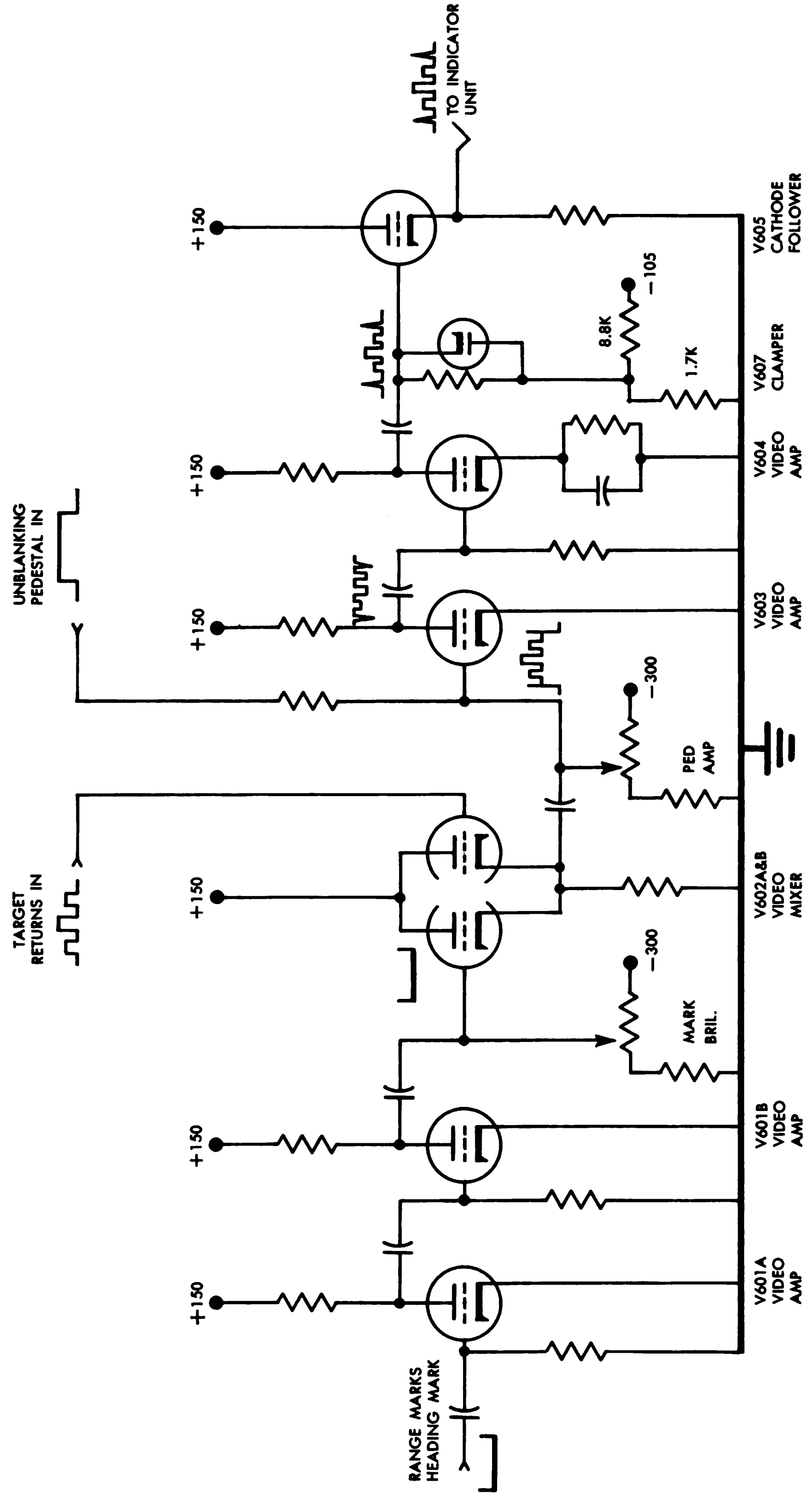
Operating With Contrast

this in the main receiver assembly schematic on page 445. For radar operation, the diode load circuit of V309 is a short time constant determined by C301 and R302 in channel one, and C302 and R304 in channel two. For pulse-stretching, resistors R303 and R304 are replaced by .1 meg-ohm resistors R301 and R303, by relay K302. This relay is energized automatically when the radar set is switched to BEACON operation.

ANTI-JAM. Anti-jamming protection against high-level CW jamming signals is provided in the grid circuit of the video limiter tubes V310 and V313. Normally relay K1603 is in the position shown and the time constant of the grid circuit of V310 is determined by the .01 microfarad coupling condenser and the 2.2 meg-ohm resistor R306. When relay K1603 is energized, the coupling condenser is reduced to 50 micromicrofarads (C303) and the grid resistor R305 to 6200 ohms. Identical changes occur in the grid circuit of the other video limiter V313. The new time constant forms a differentiating circuit which lets only the leading edges of the signal pulses drive the limiter grids negative; the positive half of the differentiated signals are clipped by grid limiting action. The received pulse width is now effectively shortened. With the ANTI-JAM switch in the ON position, the target returns appear as sharp instantaneous returns on the indicator. The anti-jam circuit is especially effective in locating targets in sea clutter or heavy cloud formations.

Video Circuit

The video circuit mixes and amplifies the target returns, range marks, and heading marker. As you can see in the schematic on page 449, the heading mark and range marks come into the circuit at the grid of V601A. They are amplified in V601A and V601B and applied to one-half of the video mixer V602. The intensity of these marks is controlled by the MARKER BRILLIANCE control, which is a bias control for V602A. Target returns come into the other grid of V602 from the main receiver assembly and are mixed with the range marks and heading mark. The combined signals are passed on to the grid of V603, a video amplifier. V603 is held below cut-off between sweeps by the fixed bias in its grid circuit. During sweep time, however, the unblanking pedestal from the sweep circuit is applied to the grid of V603,



Simplified Schematic of Video Circuit

bringing it up to the cut-off point and allowing the video signals to pass. V603 is effectively a gate which is open only during sweep time. Its bias is determined by the PED AMP adjustment, which is an operator adjustment.

The signals receive further amplification in V604 and are then applied to cathode follower V605. In the input circuit of V605 is a positive clamper biased at −17 volts. Its purpose is twofold: (1) to prevent distortion of the video signals and (2) insure that all target returns have a common reference point.

The video signals are taken across the cathode resistor of V605 and passed on to the indicator unit. V605 matches the high impedance of the video circuit to the low impedance of the coaxial cable.

Automatic Frequency Control Circuit

Most airborne radar search sets transmit at a frequency of 9375 megacycles. The radar local oscillator is usually tuned 60 megacycles below this frequency, producing a 60-megacycle IF. However, the magnetron, which is almost always used as the transmitter, is not a very stable circuit. With changes in temperature and pressure, the magnetron may shift frequency as much as 40 or 50 megacycles. Obviously, if the local oscillator is tuned to 9315 megacycles and the magnetron shifts frequency, the IF will not be the desired 60 megacycles. For this reason, radar sets have an automatic frequency control circuit (AFC). When the magnetron shifts its frequency, the AFC circuit automatically changes the frequency of the local oscillator concerned so that the IF will always be 60 megacycles.

The search AFC circuit, shown in the illustration to the right, consists of three IF amplifiers V501, V502, and V503; a discriminator detector tube V504 which develops a DC output voltage proportional to the deviation of the IF signal from 60 megacycles; a discriminator amplifier tube V505; and an AFC sweep generator tube V507 with its diode control tube V506. When the AFC-MAN switch in the radar control unit is in the AFC position, supply voltage is applied through V507 to the search local oscillator repeller plate. In the MAN position, the sweep generator is by-passed and the AFC is inoperative.

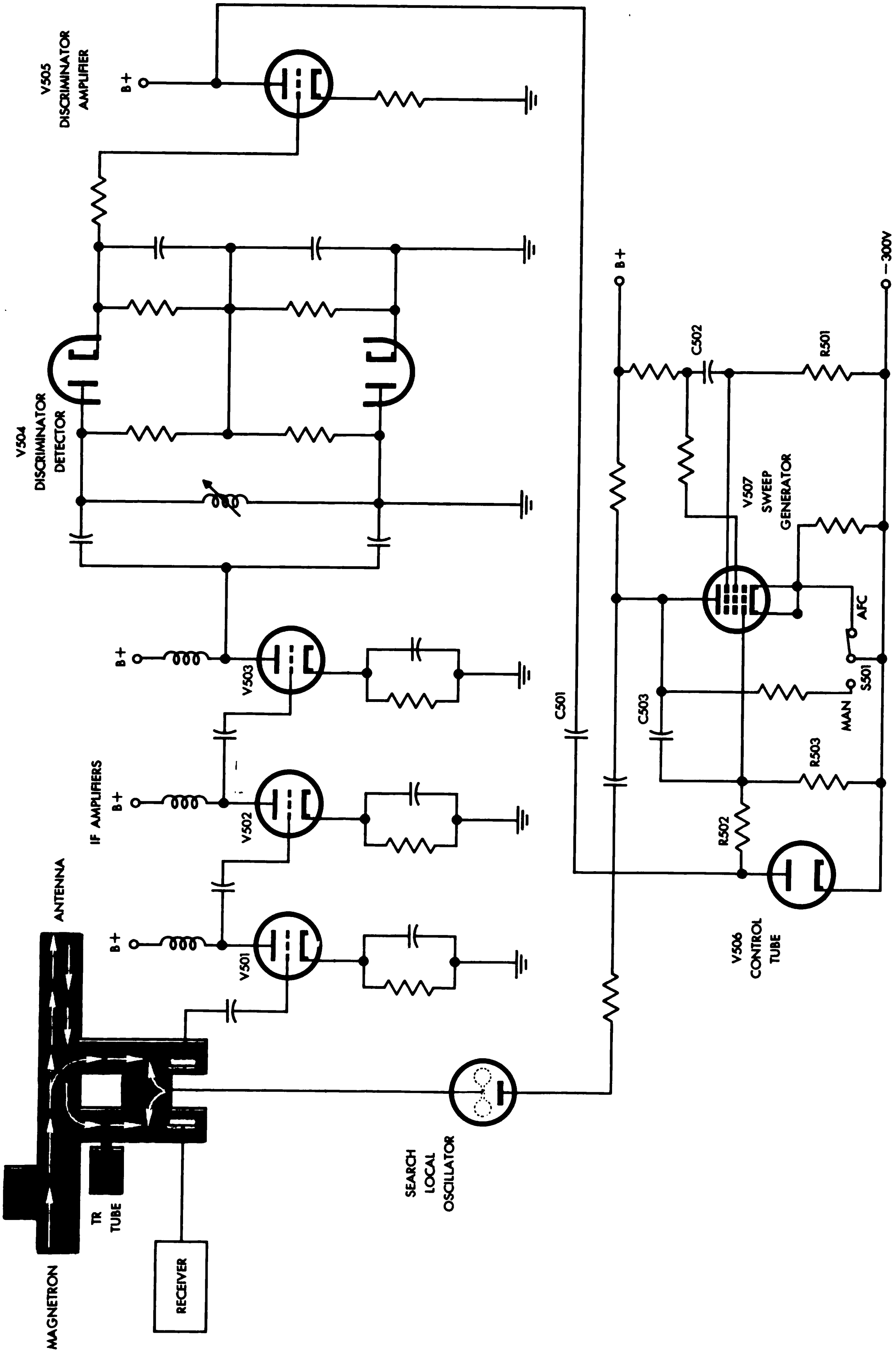
AFC IF AMPLIFIERS. The IF amplifiers are single-tuned circuits which are factory tuned to 60 megacycles. They amplify a portion of the rectified output of the receiver crystal CR-1 and apply this amplified 60-megacycle signal to the discriminator detector stage.

DISCRIMINATOR DETECTOR. If the beat frequency between magnetron and search local oscillator is below 60 megacycles, the output of the discriminator is a negative pulse each time the magnetron fires. If the beat frequency is above 60 megacycles, the discriminator puts out positive pulses. At exactly 60 megacycles, the output is zero.

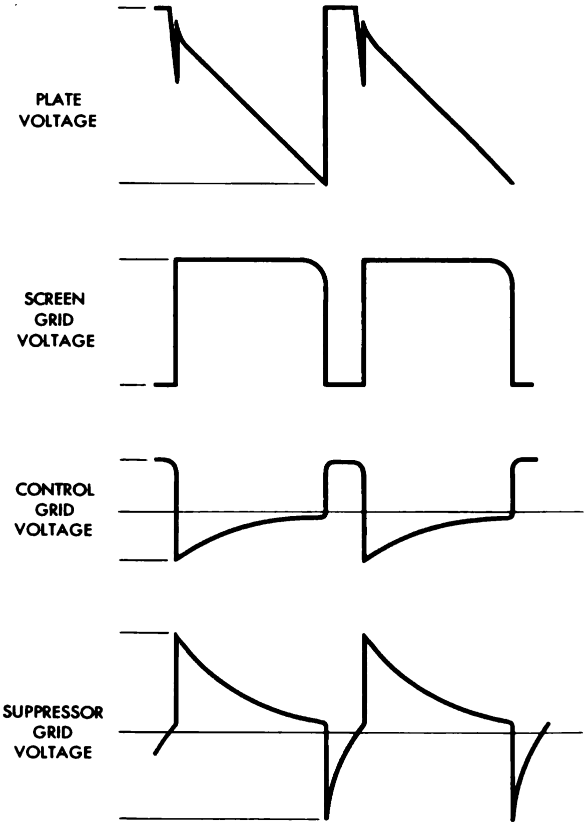
DISCRIMINATOR AMPLIFIER. The output pulses of the discriminator-detector circuit are direct-coupled to the discriminator amplifier, V505. Here they are amplified and inverted. Thus the negative pulses from the discriminator are inverted and become positive pulses in the discriminator amplifier output, and the positive pulses from the discriminator become negative amplified pulses.

AFC SWEEP GENERATOR. The AFC sweep circuit consists of a sweep generator tube V507 which generates a slow sweep to vary the local oscillator frequency over a wide range when free-running, and a sweep control tube V506 which locks the sweep on positive pulses. When the local oscillator frequency, beating with the incoming signal, causes the IF to decrease below 60 megacycles, positive output pulses from the discriminator amplifier (V505) lock in the sweep generator at the pulse repetition rate and limit the local oscillator frequency to a low value. Only a small variation about the correct IF is produced. Negative pulses do not affect the free-running sweep. Actual locking-in of the sweep generator is done by the diode control tube V506. The positive pulses draw pulses of diode current, building up a negative charge on capacitor C501 which biases and locks in the AFC sweep generator.

Free-Running Sweep. As you read the discussion which follows, study the waveshapes shown on page 452. Assume that the cycle is at the point where the plate voltage of the sweep generator starts to fall. At this time, the suppressor grid has a high positive potential because a positive voltage, developed by the previous



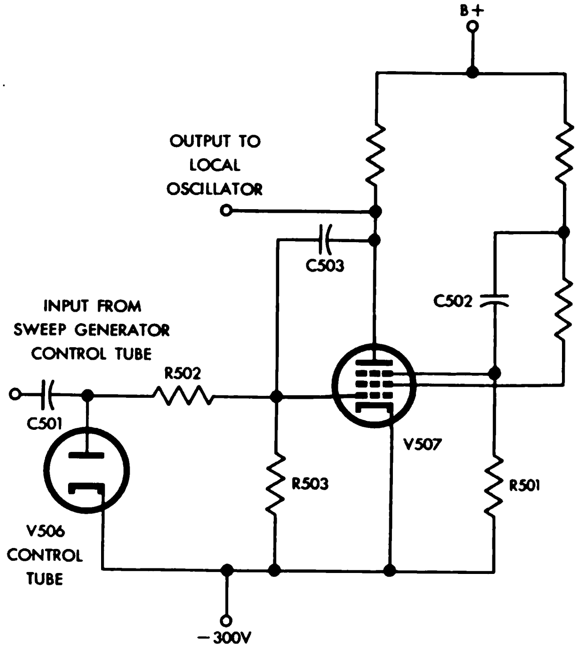
Search AFC Circuit — Simplified Schematic sharp rise in screen voltage is coupled through capacitor C502. This positive suppressor potential starts the increase in plate current and the consequent fall of plate voltage. This decline is coupled through capacitor C503 to the control grid which gradually becomes less negative. As the plate voltage reaches minimum, the control grid potential changes direction and rises sharply in a positive direction. Due to the low potential at both the suppressor and plate at this time, the control grid voltage change causes a rapid increase in screen current. The resulting drop in screen voltage, coupled through capacitor C502 to the suppressor, drives the suppressor grid sharply negative and cuts off the plate current. The plate voltage consequently rises sharply to maximum; this rise, coupled through capacitor C503, drives the control grid positive and the screen current to saturation. The screen current and the plate voltage remain at maximum while the charge on capacitor C502 leaks off through resistor R501. As the suppressor potential rises in a positive direction above a certain critical value, a pulse of plate current occurs. The resulting spike of plate voltage drives the control grid negative, and causes the screen current to be cut off, returning the screen potential to the maximum positive value. This negative spike of plate voltage occurs as the circuit *flops*, transferring control action of the control grid from the screen back to the plate. Following this change, the plate voltage starts to decline linearly again and the entire cycle is repeated.



Waveshapes of Sweep Generator Tube

Locked-in Sweep. The AFC sweep generator V507 is redrawn in the illustration to the right to emphasize the control action of the control grid and the suppressor, as well as the electron-collecting action of the screen and the plate.

The AFC control action is established in the following manner. When the free-running AFC sweep voltage starts to decline, the local oscillator frequency increases. As the beat frequency comes within the range of the discriminator bandpass, positive pulses are developed by discriminator tube V504. By inversion, these pulses appear as negative pulses at the plate of the discriminator amplifier tube V505 and are applied to the control grid of tube V507. Because their amplitude is low, these negative pulses do not affect the AFC sweep, and the


B+
OUTPUT TO LOCAL OSCILLATOR
C503
C502
INPUT FROM SWEEP GENERATOR CONTROL TUBE
R502
C501
V507
R503
R501
V506 CONTROL TUBE
-300V

AFC Sweep Generator and Control Tube

local oscillator repeller plate voltage continues to decline. When the beat frequency goes past the 60-megacycle crossover point, the polarity of the discriminator pulse output is reversed, and positive pulses appear at the output of tube V505. As each positive pulse is applied to capacitor, C501, charging current flows rapidly through the low resistance path provided by the control tube V506. When the peak of the pulse passes, tube V506 ceases to conduct and capacitor C501 discharges slowly through resistors R502 and R503. The decline of the AFC sweep voltage is so slow that as each pulse occurs, it is integrated with the preceding pulses. Since the discriminator response rises sharply from the crossover (60 megacycle) point, each succeeding pulse increases in amplitude. The bias becomes negative enough to cut off the sweep after several pulses. The sweep then reverses direction and tries to rise rapidly to a peak value. However, this means that the IF starts to increase, and the positive voltage output of the discriminator amplifier begins to lose amplitude falling to zero at the crossover point. Losing its negative bias voltage, the sweep generator once more changes the sweep direction and resumes its original slow decline until the positive pulses from the discriminator amplifier again cut off the sweep. When locked in by this sequence of functions, the local oscillator output is varied at approximately the correct frequency relationship with the magnetron, keeping the IF beat frequency varying about the discriminator crossover point.

AFC Sweep Rate. For short-range, high definition operation, the receiver IF bandwidth is 8 megacycles. The local oscillator frequency therefore can be varied over a relatively wide range without loss of signal level. Consequently, the AFC does not have to be very accurate. However, in long range search operation, the IF bandwidth is 3 megacycles. It is therefore necessary to restrict the frequency sweep of the locked-in AFC. The frequency sweep of the local oscillator with the AFC locked-in is restricted by limiting the amplitude of the voltage applied to the repeller plate of the local oscillator, thus producing only a small frequency variation. This amplitude is controlled by the rate or slope of the sweep and by the pulse repetition rate. By decreasing the sweep rate (reducing the steepness of the sweep waveshape), there is a smaller change in voltage in a given time. Since the voltage continues to fall during the interval between pulses, and the pulse rate is reduced to 200 pps on long range, the frequency of the free-running sweep has to be reduced to ½ cps to insure only a small change in sweep voltage and frequency between pulses. The transfer from fast to slow sweep occurs automatically when the radar set is switched into long-range operation.

RADAR INDICATOR CHANNEL

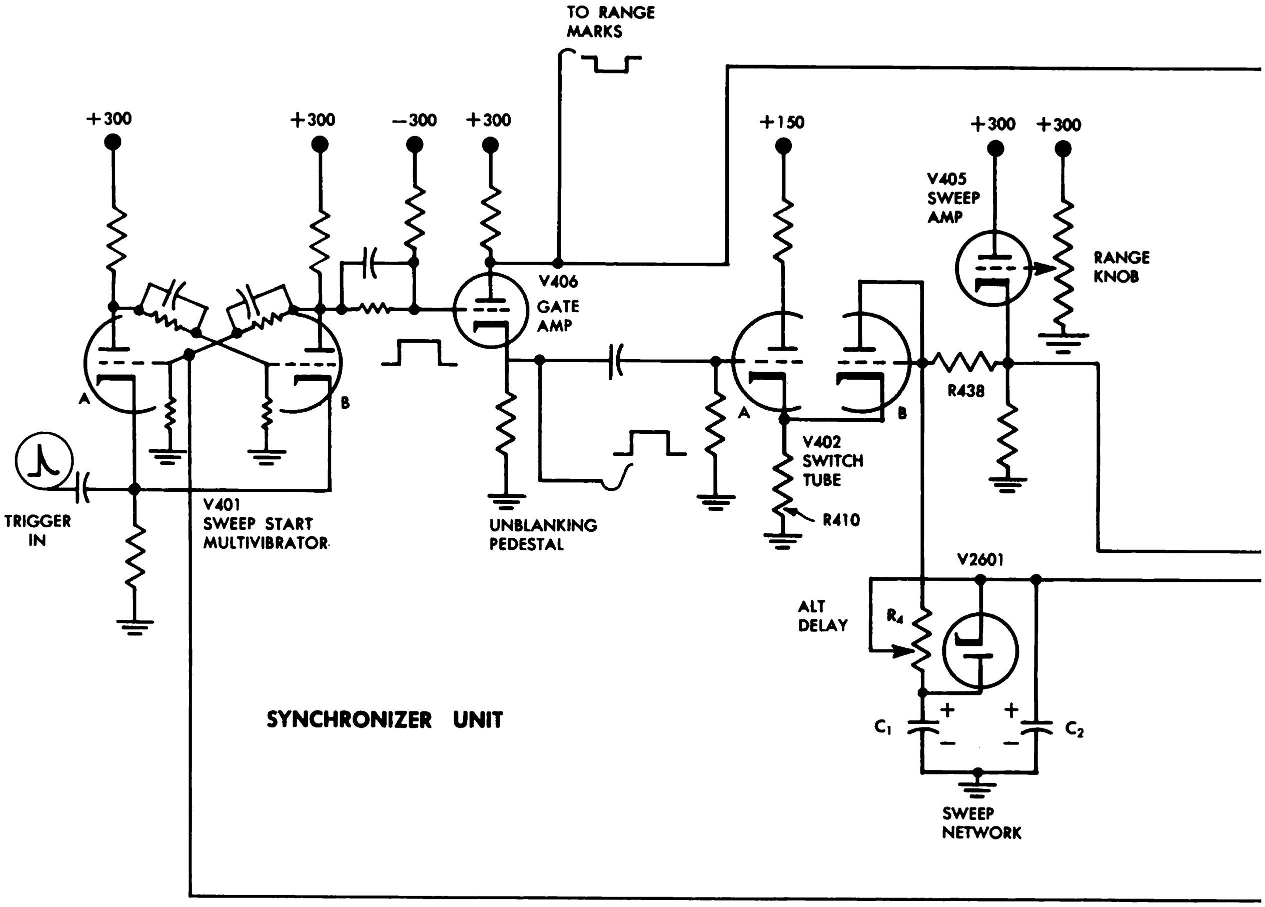
Having studied the transmitter and receiver channels, you are now ready to connect them to the third and last channel of a radar system, the indicator channel. You have already studied the circuits necessary to construct an indicator channel. Now you will see how the circuits are placed in the proper sequence to deflect the electron beam across the face of the scope and display target returns on the scope.

A Basic Indicator Circuit

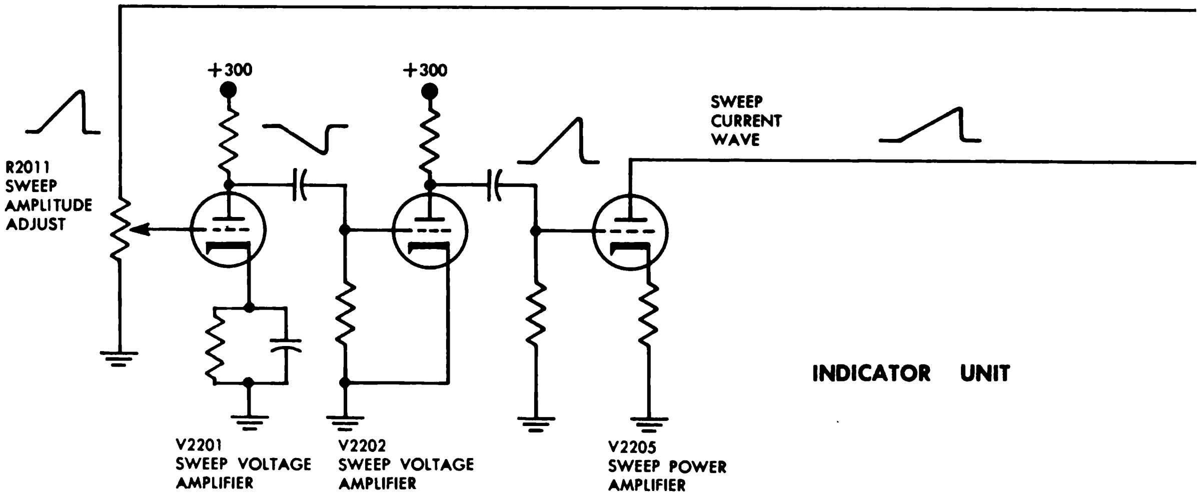
The illustration on page 458 shows the block diagram of a basic indicator channel. The trigger pulses coming from the timer are applied to the transmitter, also, therefore the sweep starts across the scope at the same instant the transmitter fires.

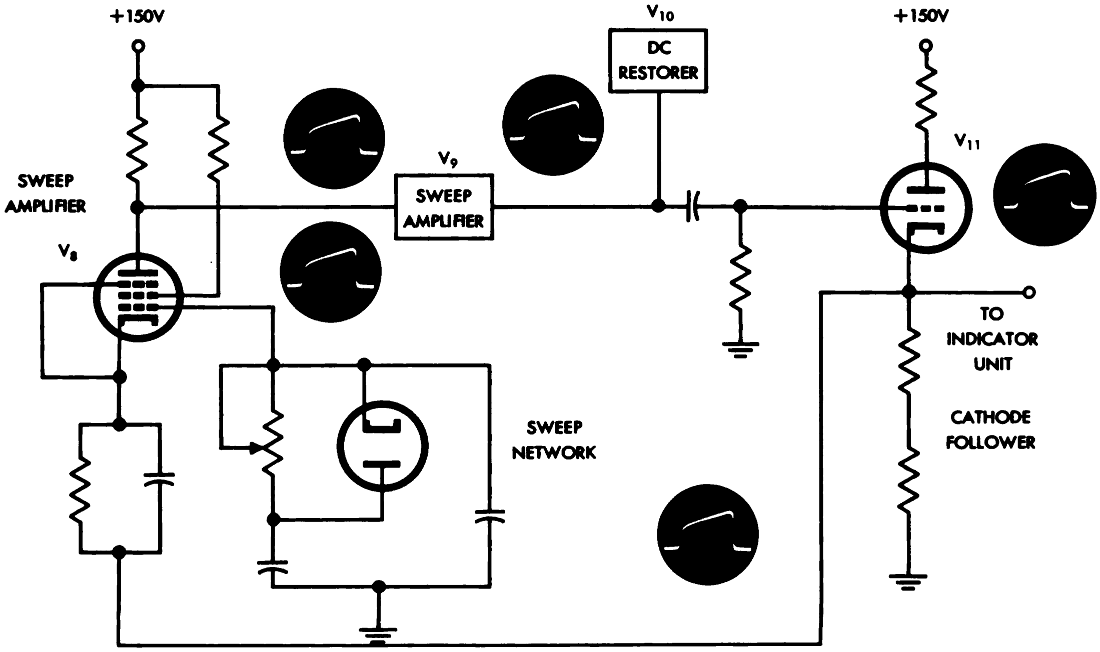
The trigger pulse trips the sweep-start multivibrator, causing it to produce a negative going square wave. This negative square wave cuts off the sweep generator, causing a capacitor to begin charging, and producing a sawtooth waveshape. This sawtooth wave is amplified by sweep amplifiers and applied as a sawtooth current waveshape to the deflection coils of the electromagnetic CRT shown. This sawtooth variation of current causes the deflection of the beam across the scope. When the sweep has built up for a predetermined period of time (corresponding to the range desired), a sweep-stop tube is tripped. Its output is then used to cut off the sweep-start multivibrator, thus ending the sweep.

In order to get the amplitude of the sawtooth waveshape great enough to deflect the electron beam across the scope, as many as seven or eight sweep amplifiers may be used. Usually two or three of these are located just following the sweep generator and the remainder in the



APS-23 INDICATOR CHANNEL


+300
+300
R2011 SWEEP AMPLITUDE ADJUST
SWEEP CURRENT WAVE
INDICATOR UNIT
V2201 SWEEP VOLTAGE AMPLIFIER
V2202 SWEEP VOLTAGE AMPLIFIER
V2205 SWEEP POWER AMPLIFIER



Producing Trapezoidal Waveshape

potential on the cathode V405. The voltage across the capacitor rises and this rising voltage is coupled through two stages of amplification, V403 and V404A, to the cathode followers, V408 and V409, in parallel. As the grid of the cathode follower rises in potential, so also does the cathode. This rising potential is coupled back to the cathode of V405. As you can see from the diagram on page 461, the capacitors in the Sweep Network began to charge from three volts toward 50 volts, a distance of 47 volts, but as soon as the capacitors began to charge, the potential toward which they charged, began to rise. So, the capacitors in the Sweep Network charge continuously toward a voltage some 47 volts higher than the instantaneous voltage across them.

Should you care to charge the capacitor more rapidly, you need only to change the setting of the Range Knob to a shorter range which increases the voltage on the grid of V405, the cathode of V405 and, hence, the voltage toward which the capacitor is to charge. The output of the cathode followers V408 and V409 is distributed to the sweep amplifiers in the indicator unit and to the grid of the Sweep Stop Tube, V404B. V404B is normally cut off by fixed bias, the amplitude of which is determined by the setting of R924 (Sweep Length Adjust). The bias on the stop tube determines the voltage to which the sweep rises before the tube conducts.

The length of time it takes the sweep to reach this voltage corresponds to the desired range. As V404B starts conducting, its plate potential drops and this voltage drop is coupled back to the left-hand grid of V401. The left side of V401 is cut off and the right-hand side comes back into conduction, bringing the sweep-start multivibrator back to normal and ending one sweep cycle.

The DC Restorer, V407, maintains a constant bias on the grids of V408 and V409, which is sufficient to keep the cathode followers cut off at all times except when a sweep is being produced. V407 is simply a biased clamper that functions similarly to the clampers previously studied.

The sweep voltage from the cathode follower is applied to the Indicator Unit where it is amplified by two voltage amplifiers, V2201 and

V2202. The sawtooth of voltage is then coupled to the two sweep power amplifiers, V2204 and V2205, in parallel where the sawtooth of voltage is converted to a sawtooth of current. There are two feedback paths from the Sweep Drivers V2204 and V2205 to the Sweep Amplifier V2201 to insure that the current in the deflection coils matches as closely as possible the voltage input to V2201. In the Indicator Unit there is also another DC Restorer, V2203. The purpose of this DCR is to prevent grid leak biasing of the Sweep Drivers and to insure that no current flows through the Sweep Coils except when a Sweep is being produced. When current flows in the Sweep Coils, the electron beam cannot be centered on the cathode ray tube. Therefore, it is imperative that no current flows in the Sweep Coils between sweeps.

THE INDICATOR. The indicator is a flat-faced cathode-ray tube with a fluorescent screen of long persistence. This long persistance is necessary to retain the image on the screen for a map-like presentation. To reduce strain on the operator's eyes, an amber filter is placed in front of the tube which reduces the brilliance of the flash resulting from electrons striking the screen.

As you can see in the illustration on page 457, a 5000 volt DC potential is used on the aquadag coating on the inner surface of the tube. The aquadag serves as a second anode and supplies a return path to ground for the electrons. A positive 300 volt DC potential is applied to the first, or accelerating, anode.

The potentiometer R2209 in the control grid circuit controls the bias on the CRT, thus controlling the intensity of the electron beam.

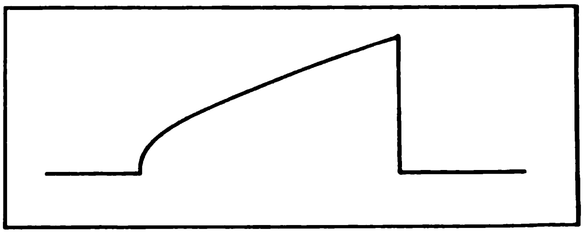
Video signals are fed into the indicator unit from the receiver through potentiometer R2010, the VIDEO GAIN control. The setting of R2010 therefore determines the intensity of the video signals on the CRT. The video signals are applied as negative-going signals to the cathode of the CRT and cause conduction through the CRT to increase. The amplitude of each video signal determines how much conduction is increased and, consequently, the intensity of each video return on the scope. Small target returns therefore do not create as bright a spot on the scope as do large target returns. This variation of intensity with strength of target return is known as *intensity modulation.*

Effect of Altitude Delay on Sweep Circuit Operation

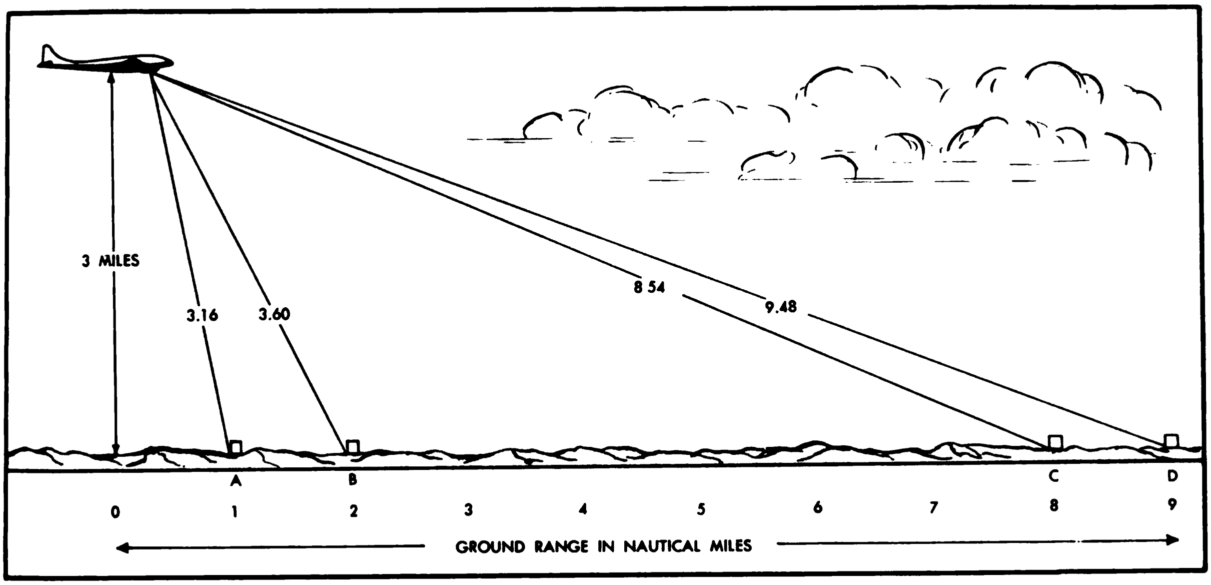
In an airborne search radar set, reflected target returns appear at a distance on the scope that is representative of their slant range from the aircraft. There is an inherent distortion of nearby targets in this type of presentation as shown in the diagram on page 461.

In this situation, the aircraft is flying at an altitude of three nautical miles, and there are two targets, A and B, spaced one mile apart on the ground. Targets C and D are also one mile apart at a ground range of 8 and 9 miles, respectively. In this situation, you would naturally expect some error because you know that you are going to read slant range to the targets on your scope, but you would probably expect to see the target returns painted in their proper perspective on the scope, i.e., target B one mile from target A, and target D one mile from target C. Note in the diagram, however, that the slant range distance to target B is only .44 nautical miles greater than the slant range to target A. Target A will be painted on the scope at a range of 3.16 nautical miles, and target B will be painted .44 nautical miles farther out, definitely an erroneous presentation. Note also that the difference in slant range between targets C and D is .94 miles, a negligible error. This distortion is appreciable only for close-in targets.

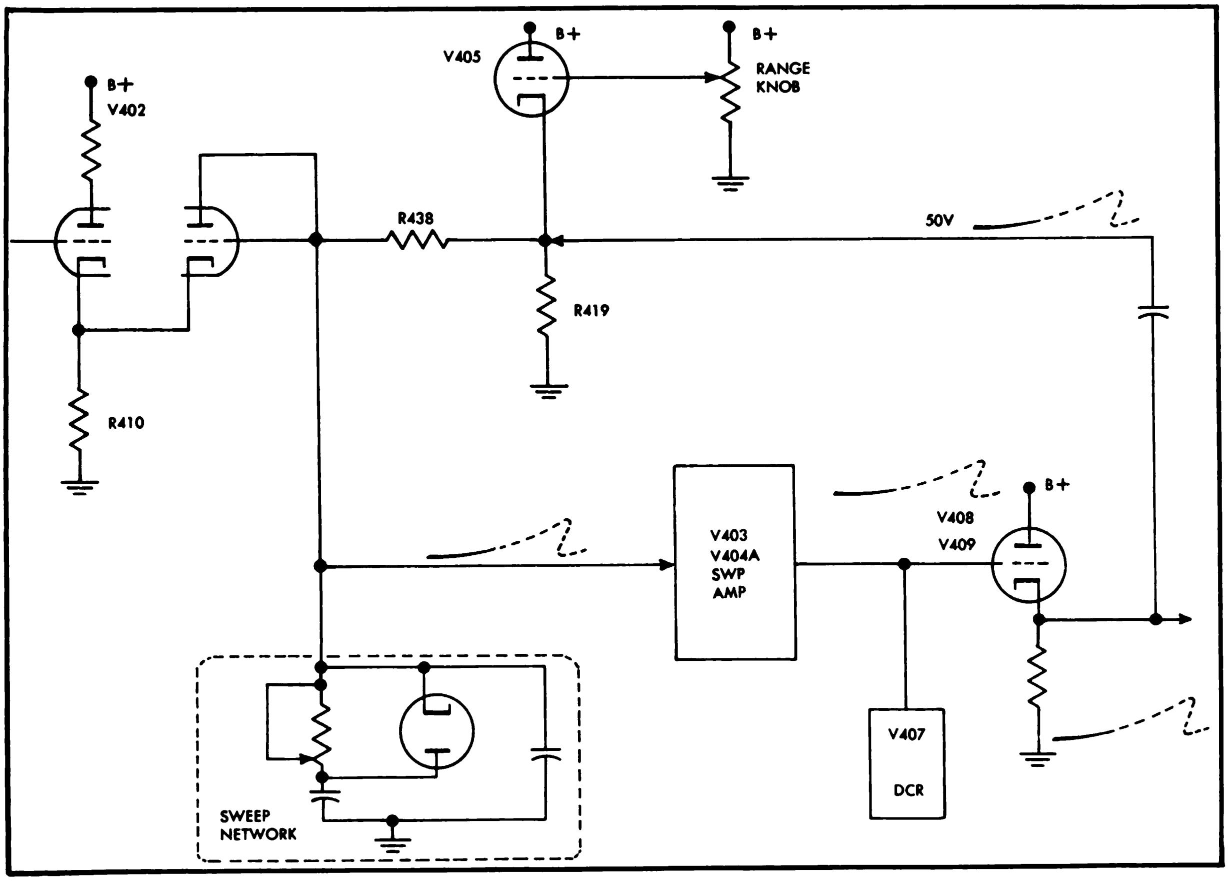
In order to prevent this distortion, and to present the targets on the scope in their proper perspective, a hyperbolic sweep is incorporated into the APS-23. This sweep waveshape is controlled by the Altitude Delay Knob, since the



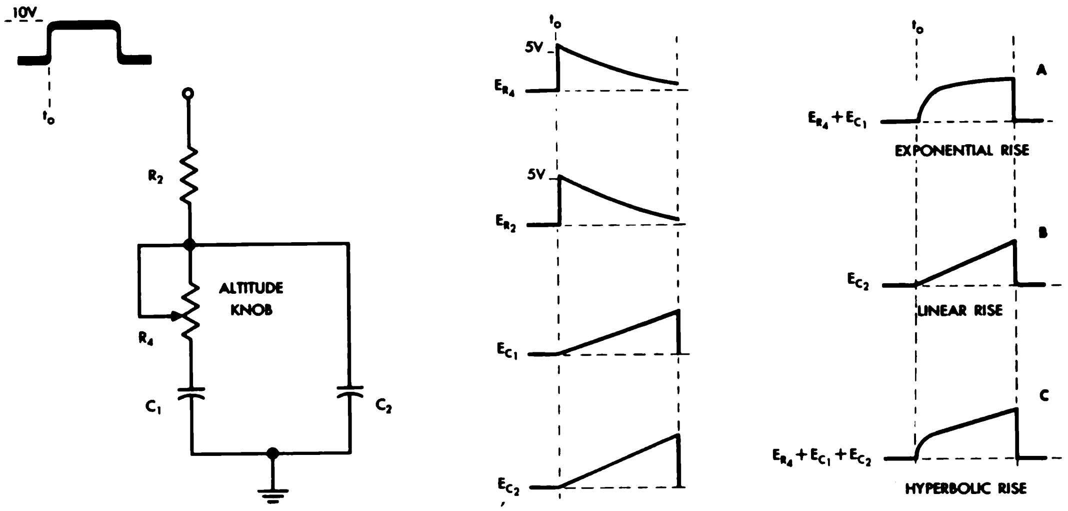
Hyperbolic Sweep



Slant Range versus Ground Range



Sweep Forming Network with Altitude Delay Set In



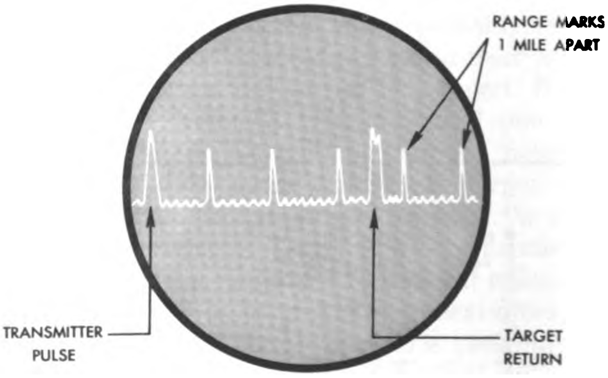
Equivalent Circuit of Sweep Network and Waveshapes Produced

error in presentation varies with the altitude of the aircraft. With no altitude delay set in, the normal sawtooth waveshape is generated. A hyperbolic wave is required, because in order to spread out nearby targets, a sweep is needed which starts out fast, and then slows down as the range increases. This type of waveshape is shown in the illustration on page 460.

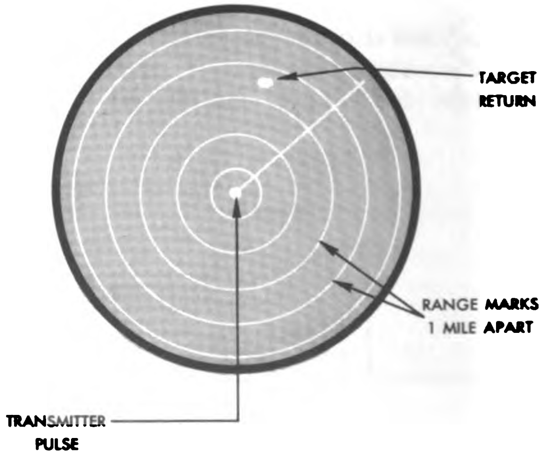
An equivalent circuit of the sweep network with waveshapes showing how the hyperbolic sweep is formed is shown on this page. C_1 has in its charge path R_4 and R_2, while C_2 has only R_2 to charge through. If you assume that a 10 volt square wave is applied to the circuit, all the voltage will be dropped across R_4 and R_2 the first instant. As the capacitors charge, the voltage across the resistors will decrease as shown. Combining the voltage across R_4 and C_1, an exponential curve is obtained as shown at A in the illustration. When this waveshape is combined with the voltage across C_2 (shown at B), a hyperbolic curve is obtained as shown at C. The purpose of V2601 (See Diagram page 457) which is in parallel with R_4 during the time C_1 is discharging, is to create a fast discharge path for C_1 when the sweep ends. During sweep time, the cathode of this tube is more positive than the plate, so the tube has no effect on circuit operation.

AUXILIARY CIRCUITS

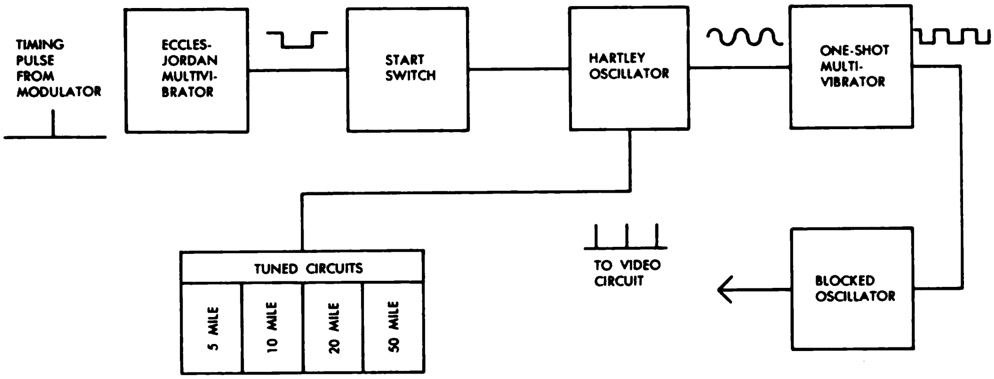
Each radar system, in addition to the three



Range Markers on A-Type Scan



Range Marker on PPI-Type Scan



Block Diagram of Range Marker Circuit

channels of which it consists, also contains auxiliary circuits that make its operation more effective. Two types of auxiliary circuits are the *range marker* and *delay circuits*.

Range Marker Circuit

In order to make estimations of range to targets more accurate, radar sets are provided with range markers. These range markers appear along the time base (sweep) at given distances apart. The distance between range markers is such that the range is divided into smaller units, and estimations of range are thus made more accurate. The illustration at the left shows how range markers appear on an A-type scan. The scope shown represents operation on a 5-mile range with range markers 1 mile apart. The target shown is approximately 3.4 miles away.

On a PPI-type scan, the range markers cause circles to appear around the face of the scope at given distances from the center. The appearance of the PPI-type scan with range marks is shown in the illustration to the left. Here also the range marks serve the purpose of making estimates of distance to targets much more accurate.

In order that range markers may serve the purpose of accurately calibrating a radar scope for measuring distances to a target, two conditions must be met:

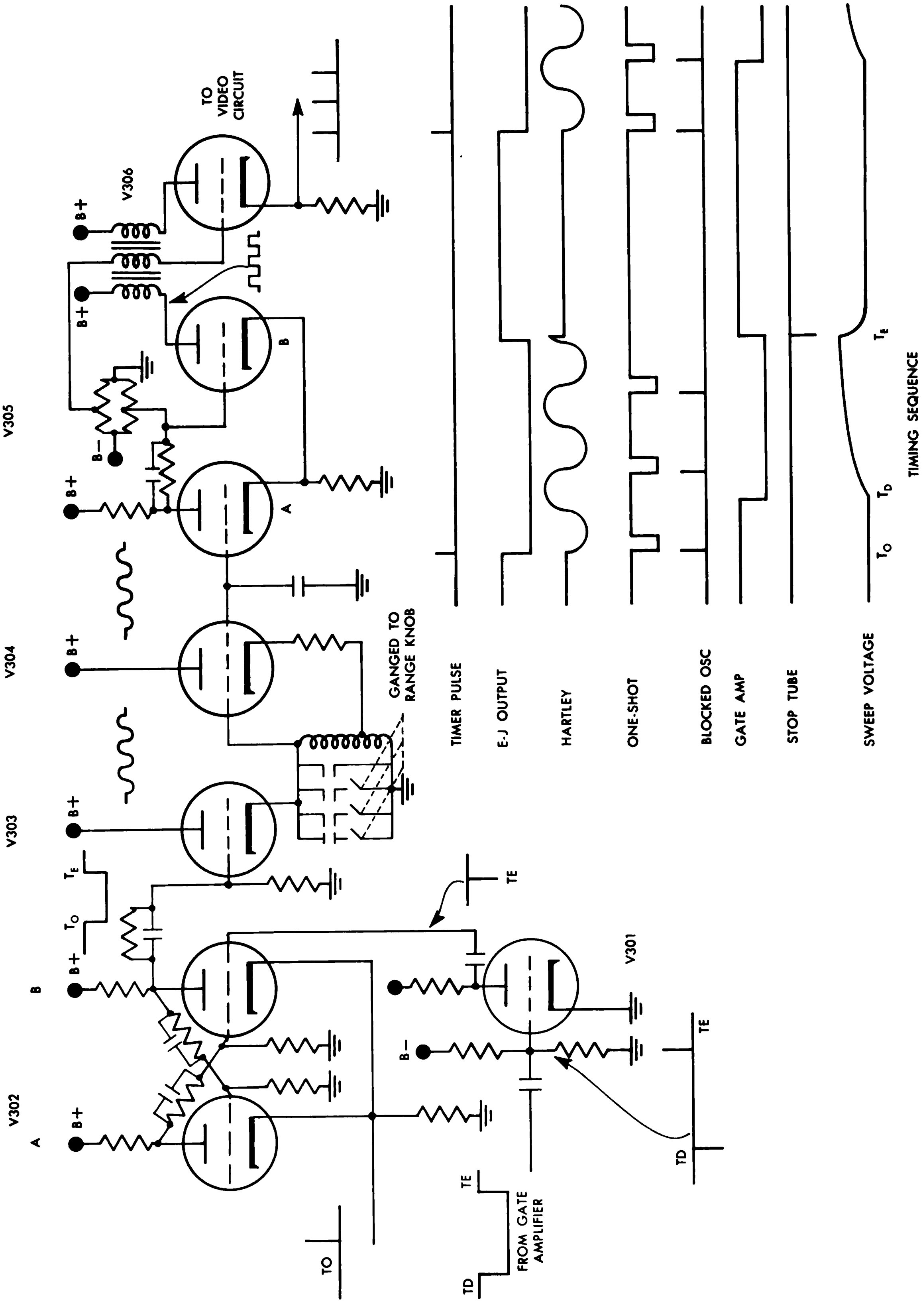
(1) The time (distances) between the range marks must be accurately determined.

(2) The range marks must be synchronized with the transmitter and sweep pulses.

TYPICAL RANGE MARKER CIRCUIT. The block diagram above shows the components necessary to form a series of narrow pulses suitable for use as range markers. The Hartley oscillator has four precision-tuned tank circuits which set the duration of each oscillation at a time equivalent to 5, 10, 20, or 50 nautical miles. Only one tuned circuit is switched in at a time, according to the range on which the radar set is operating. The positive peak of each sine wave from the Hartley oscillator triggers the blocked oscillator into conduction, producing very narrow pulses which are sent on to the video circuit and eventually to the CRT where they are displayed as range marks.

APS-23 RANGE MARKER CIRCUIT. The block diagram above shows the use of a Hartley Oscillator and a blocked oscillator to form a series of narrow pulses suitable for use as range markers. The timing pulse from the Thyratron in the modulator of the transmitter channel triggers the Eccles-Jordan V-302 multivibrator, which begins to produce a negative-going rectangular wave. This rectangular wave cuts off the start switch V303 and allows the Hartley oscillator V304 to begin oscillating. The oscillations from the Hartley oscillator are coupled to a one-shot multivibrator V305 where a short duration rectangular wave is produced for each cycle of sine wave input. The short duration rectangular wave triggers the blocked oscillator, which produces a narrow positive pip voltage that is sent to the video mixer where it is added to the target returns. Eventually these marks arrive at the cathode of the CRT where they are displayed as range marks.

CIRCUIT ANALYSIS. The illustration following is a simplified schematic of the APS-23 range



Simplified Schematic of APS-23 Range Marker Generator marker generator. It is the same circuit that was discussed in block diagram on page 463. Before the trigger arrives from the transmitter channel, the start switch V303 is conducting heavily. The conduction causes a large amount of screen grid (not shown) current and, hence, a low screen grid voltage at V303. The screen of V304 is connected to the screen grid of V303 and the low screen grid voltage prohibits the Hartley oscillator tube V304 from amplifying sufficiently to allow oscillations. The One-Shot multivibrator is a bit different than those discussed earlier in this manual. V305B is biased just below cut-off by the bleeder network and V305A is conducting. As soon as the input to V305A drops slightly in potential, the tube (V305A) is cut off, V305B conducts, and a narrow negative-going rectangular wave is formed.

When the timer pulse, at time zero, arrives at the common cathode of the Eccles-Jordan multivibrator, the left-hand plate cuts off. Its rising potential brings the right-hand plate into conduction and the beginning of a negative-going rectangular wave is produced. This negative-going potential is coupled to the grid of the start switch V303, cutting it off. When V303 cuts off, the Hartley oscillator begins oscillating at a frequency equivalent to the Range Marks desired. The oscillations are coupled out of the Hartley tank to the grid of V305A. As soon as the potential across the tank drops or goes in a negative direction from its quiescent voltage, V305A is cut off allowing V305B to begin conducting. Each sine wave from the Hartley oscillator produces one negative-going rectangular wave at the plate of V305B. These pulses are coupled through transformer T301 with such polarity that the leading edge of each rectangular wave brings the blocked oscillator into conduction producing a peaked wave at its cathode. These peaked waves are fed into the video circuit where they are amplified, mixed with the other video signals, and then applied to the CRT, where they are displayed as range marks.

When the sweep ends, one of the outputs from the Gate Amplifier V406 (a negative gate) is coupled to the Stop Tube V301. At the input to the stop tube, the gate is differentiated across a short time-constant circuit. Therefore, there will be a negative pip at the grid of V301 when the sweep starts and a positive pip when the sweep ends. The stop tube is normally cut off by fixed bias; therefore, the negative pip will have no effect on the tube. The positive pip, however, will cause the stop tube to conduct for a short period of time, producing a negative-going pip at the plate which is coupled to the grid of the conducting side of the Eccles-Jordan multivibrator. The conducting side is cut off, allowing the cut-off side to begin conducting again, the start switch comes back into conduction, and the entire circuit returns to its original condition.

The timing relationship can be seen under the schematic.

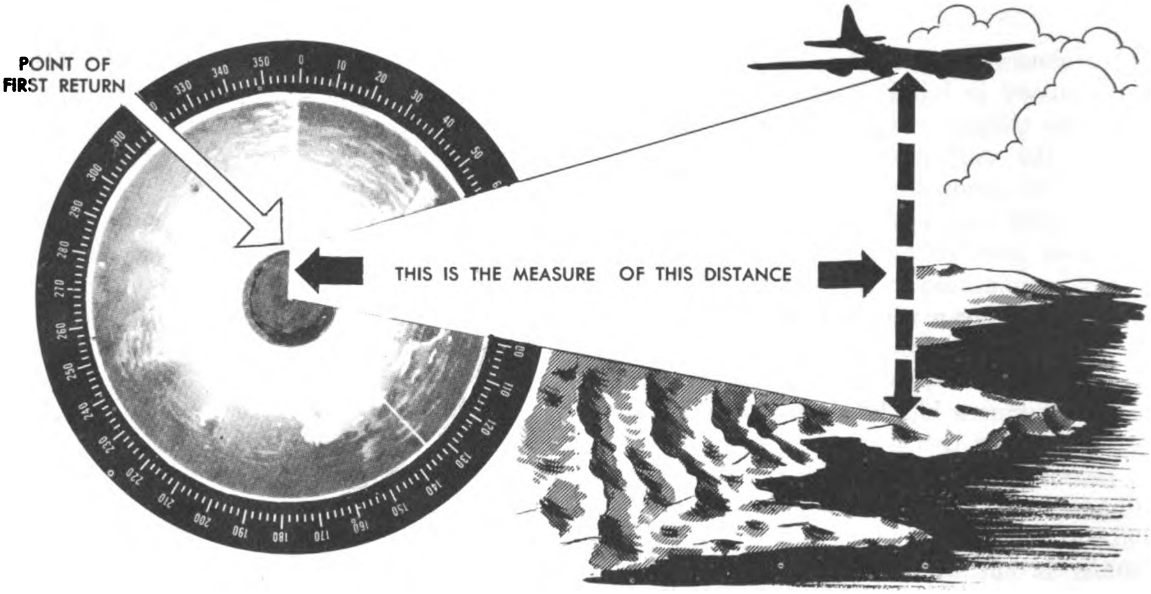
Delay Circuits

If the sweep on the CRT starts across the scope at the same instant that the transmitter fires, the first targets seen are those directly beneath the aircraft. Therefore, on the PPI scope, a hole with a radius corresponding to the altitude of the aircraft appears in the center as shown in the illustration on page 466, Need for Altitude Delay.

Objects which are directly beneath the aircraft appear on the scope at a distance equal to the distance between the aircraft and the ground.

To prevent having this area of no return on the scope, airborne radar equipment is provided with an Altitude Delay arrangement. The start of the sweep is delayed by an amount equal to the altitude of the aircraft. With altitude delay in, the hole is closed up and objects directly underneath the aircraft appear at the center of the screen. Altitude delay also performs another function. Since the delay set in measures the distance between the aircraft and the ground, it can be used as an indication of absolute altitude.

The block diagram on page 466 shows how an altitude delay circuit could be incorporated into a radar set. With the altitude delay switch in the off position, the trigger pulse coming from the timer is applied to both the transmitter and the sweep generator at the same time (t_0). Therefore the sweep starts across the CRT at the instant the transmitter fires as shown by the waveshapes.

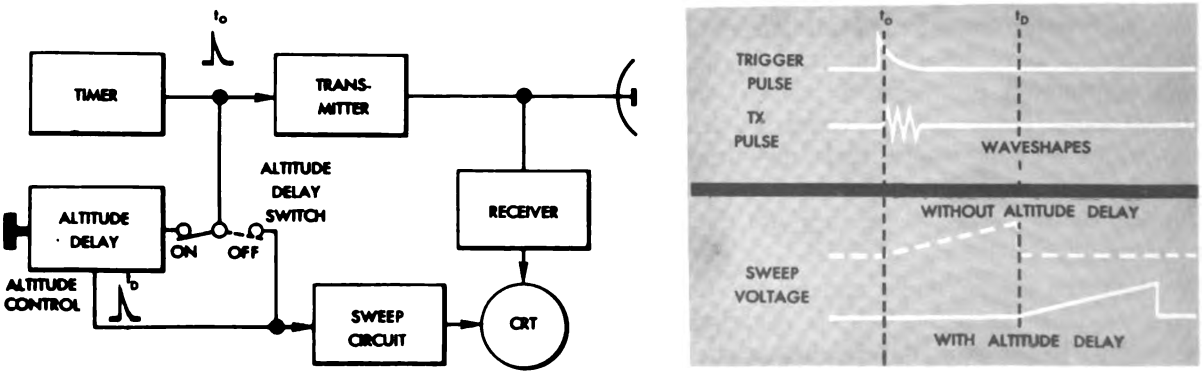


Need for Altitude Delay

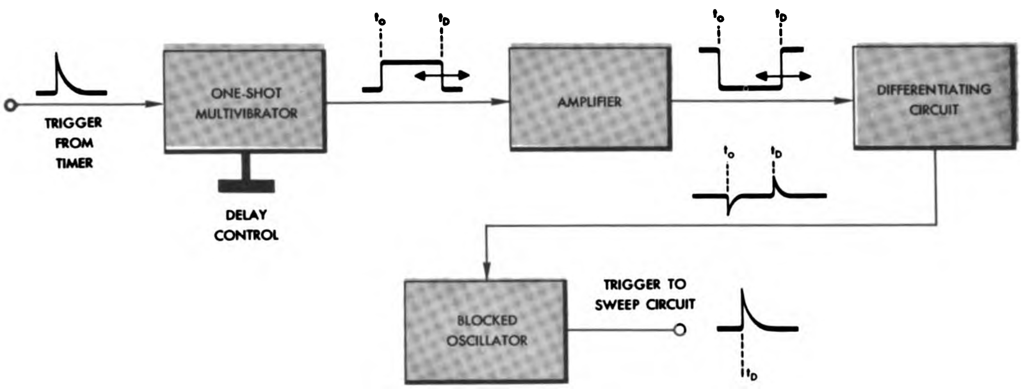
Now place altitude delay in the circuit by moving the switch to the ON position. The trigger pulse, occurring at t_0 is applied directly to the transmitter; therefore, the transmitter still fires at t_0. Before the trigger pulse can be applied to the sweep generator, however, it must pass through the altitude delay circuit. The trigger pulse is delayed in this circuit by an amount dependent upon the setting of the Altitude Delay Control (ADC). After this amount of delay (t_D), the trigger pulse goes on to the sweep generator and the sweep starts across the scope at time t_D as shown by the waveshapes. The amount of delay set in is normally equal to the altitude of the aircraft above the terrain.

BASIC DELAY CIRCUIT. Note that the block diagram of a basic delay circuit, shown at top of page 467, consists of a one-shot multivibrator, amplifier, a differentiating circuit, and a blocked oscillator. By analyzing the waveshapes, you can see that the start of the sweep can be delayed any amount desired by the Delay Control knob in the multivibrator.

The same circuit is shown in schematic form in the illustration below it. The one-shot multivibrator is triggered at time zero (t_0) by the pulse from the timer. The width of the output pulse of the multivibrator is determined by the setting of the Delay Control. As denoted by the arrow, the trailing edge of the square wave is, then, variable. Assume in this case that the square wave is 10 μs wide. This positive-going square wave is amplified and inverted by V_3 and immediately differentiated by the short time-



Block Diagram of Altitude Delay Circuit



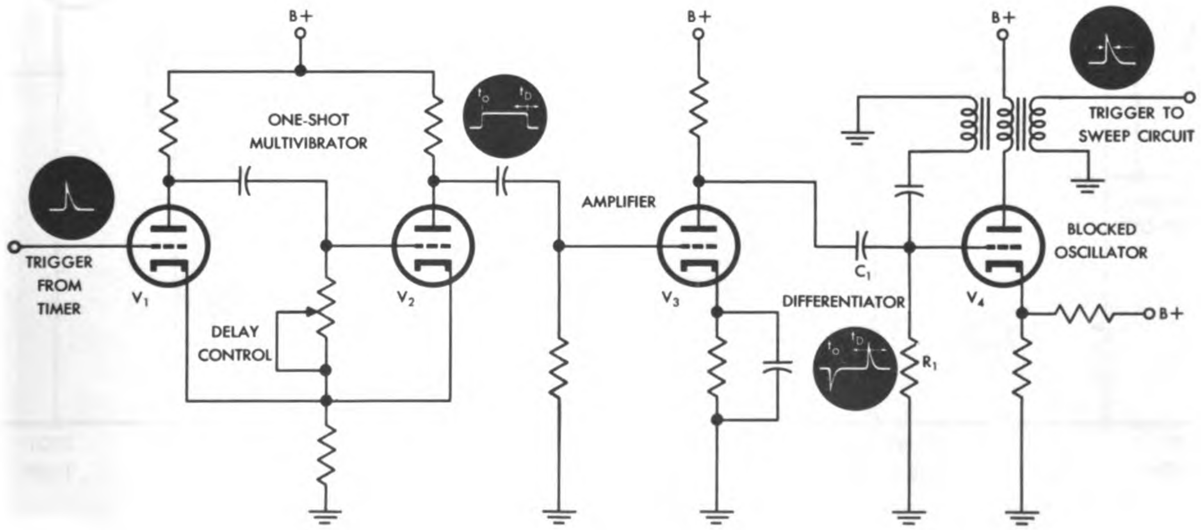
Block Diagram of Delay Circuit

constant circuit R_1 C_1. The values of R, and C, are chosen so that they form a short time-constant for the square wave. The voltage across R_1, applied to the grid of the blocked oscillator V_4, is a negative-going peak at t_0 and a positive-going peak occurring 10μs later (t_D). Since the blocked oscillator is held at cutoff by the fixed bias on its cathode, the negative peak has no effect on the tube. The positive peak, however, which occurs at time t_D triggers the blocked oscillator, and a positive-going peaked wave is produced across the tertiary winding of the transformer. The purpose of the blocked oscillator is to eliminate the negative peak and to sharpen up the positive peak. This positive-going, peaked wave then triggers the sweep circuit. The sweep therefore starts across the scope 10μs after the transmitter fires.

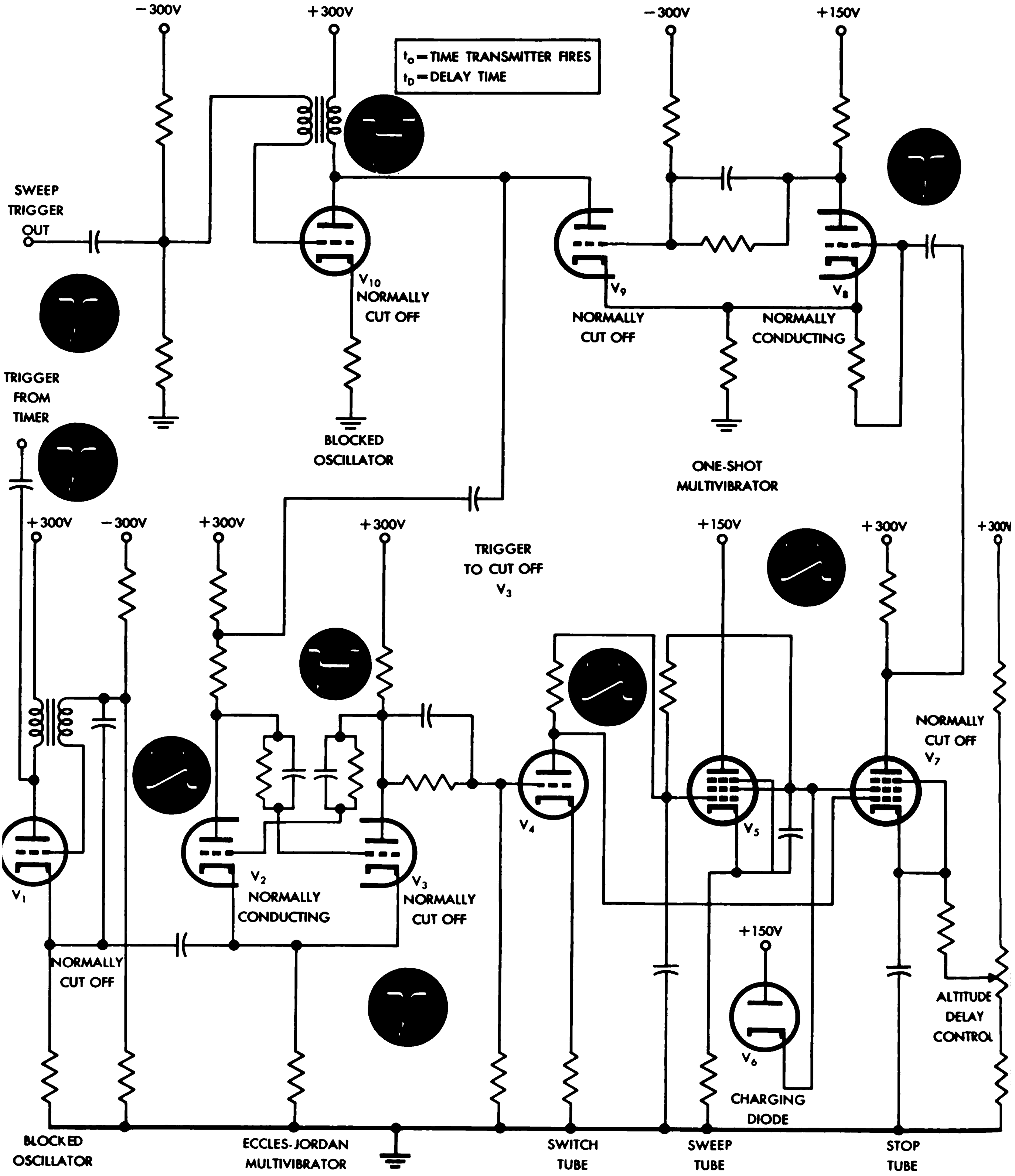
Although the above circuit is a practical circuit and will produce any amount of delay desired, it is not as precise a circuit as is required in some of our radar sets. The one-shot multivibrator is not a highly stable circuit. That is, the width of its output may vary slightly from one pulse to the next. Obviously, a circuit of this type is not desirable in a precision instrument. The remainder of this chapter is devoted to an analysis of two very accurate delay circuits, the first of which is used in the AN/APS-23 and the other in the APS-42.

AN/APS-23 DELAY CIRCUIT. Study the block diagram and schematic of the AN/APS-23 delay circuit shown on pages 468-469. The method of producing an accurate delay is as follows:

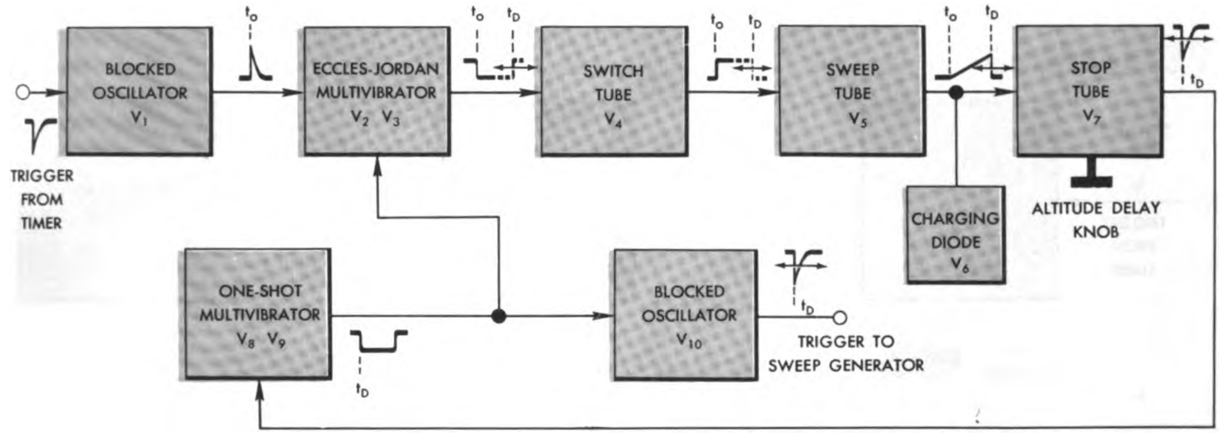
The trigger pulse from the timer fires the



Schematic of Delay Circuit



Schematic Diagram of AN/APS-23 Delay Circuit

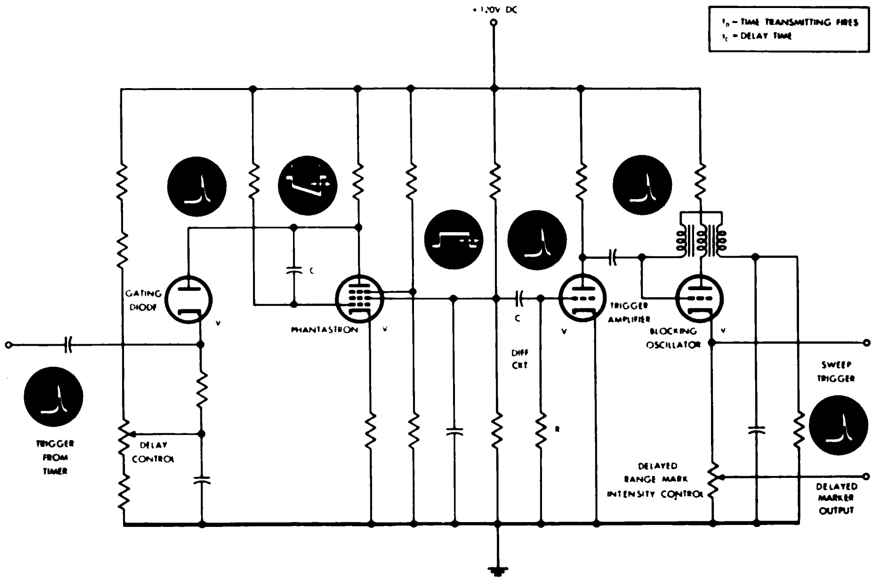


Block Diagram of AN/APS-23 Delay Circuit

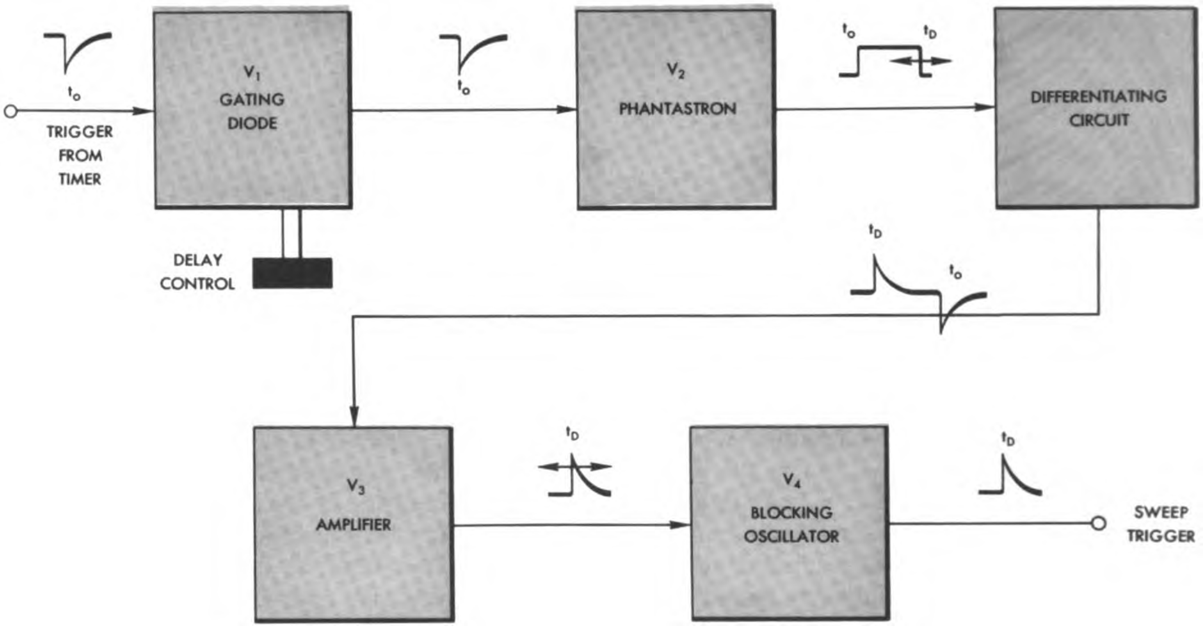
blocked oscillator (V_1), which produces a peaked pulse. This pulse is then used to trip the Eccles-Jordan multivibrator (V_2 and V_3). As right half of this multivibrator (V_3) starts conducting, its plate potential drops. The multivibrator remains with V_3 conducting until another pulse is applied to complete the cycle. This pulse will not be applied until delay time (t_D) later.

The drop in plate potential of V_3 cuts off the switch tube (V_4) which allows a condenser in the circuit of the sweep tube (V_5 and V_6) to begin charging. The stop tube (V_7) is held below cutoff by a fixed bias proportional to altitude and controlled by the Altitude Delay knob. When the condenser charges sufficiently to overcome this fixed bias, the stop tube conducts and triggers the one-shot multivibrator (V_8 and V_9). The steep leading edge of the one-shot multivibrator output cuts off the Eccles-Jordan multivibrator, ending its output. The output of the one-shot multivibrator also trips a blocked oscillator V_{10} and produces the sweep trigger. This sweep trigger occurs at time t_D, the amount of delay set in by the Altitude Delay knob.

Note that the leading edge of the stop tube output waveshape is sloping. The only function of the one-shot multivibrator is to cause this edge to be sharp so that the amount of delay will be exactly as desired. This is a very accurate circuit which produces exactly the amount of



Schematic Diagram of APS-42 Delay Circuit



Block Diagram of APS-42 Delay Circuit

delay the operator desires.

AN/APS-42 DELAY CIRCUIT. The delay circuit in the AN/APS-42 is essentially the same as the basic delay circuit studied earlier in this chapter. However, a phantastron circuit replaces the one-shot multivibrator as a square-wave generator. A phantastron may be considered to be a multivibrator (in one tube) which is more accurate and stable than the conventional multivibrator. Study the block diagram above and the schematic of the AN/APS-42 delay circuit on the previous page. The method of producing an accurate delay is as follows:

The gating diode has a fixed bias on its cathode proportional to the amount of delay desired and set by the DELAY CONTROL potentiometer. Since the plate of the gating diode and the plate of the phantastron are connected, the conduction of the diode sets the potential on the plate of the phantastron. This potential determines the amount of delay to be obtained. The conduction in the phantastron circuit is from the cathode to the screen grid before the trigger pulse is applied.

The negative trigger pulse from the timer causes the gating diode to conduct more heavily, dropping the plate potential suddenly. This sudden change in plate voltage is coupled to the control grid of the phantastron tube through C, causing less conduction to the screen grid. The leading edge of the positive output on the screen occurs at time zero (t_0) and the trailing edge is determined by the delay set in (t_D). The output from the screen grid is therefore a positive-going square wave whose duration is determined by the delay desired. At the end of this time, the plate returns to its normal potential determined by the setting of the DELAY CONTROL, until the next trigger pulse arrives from the timer.

The square wave output from the screen grid of the phantastron is differentiated by C_1 and R_1. This differentiation produces a positive peak at time t_0 and a negative peak at time t_D. The positive peak is clipped and the negative peak inverted in V_3. The output of V_3 is a positive peak occurring at time t_D and used to trigger the blocking oscillator V_4. The output of V_4 is a very narrow pulse occurring at t_D which is used to trigger the sweep circuit. Thus, the start of the sweep is delayed by a time equal to t_D.

Chapter 39

Synchros and Servomechanisms

There are many operations in which it is desirable for two shafts to rotate in synchronization. But because the distance between the two shafts is too great, or because one shaft cannot develop enough torque to rotate the other one, it is not always feasible for the shafts to be connected together mechanically. Therefore, shafts which are designed to rotate in synchronization are most generally connected electrically. Systems employing electricity for rotating two shafts in synchronization are called remote indication systems.

The *synchro* is a machine which converts mechanical position into electrical position, or electrical position into mechanical position; and the *servomechanism* is a system which amplifies and transmits mechanical position from one position to another by electrical means.

A remote-indication system is designed to produce at one position an indication of an operation at another position, or to transfer information from one position to another. Sometimes this system is called a data transmission system. Many radar antennas, for example, are designed to rotate 360° but are located in places where it is not possible for the operator to see them. Therefore, a data transmission system is employed to give the operator the direction in which the antenna is pointed. A pointer attached to a shaft rotates in synchronization with a shaft connected to the antenna, and registers the position of the antenna on a dial at the operator's position. The devices through which the position of the antenna is transmitted to the dial are called synchros. They resemble small electric motors. The synchro at the antenna is called a synchro generator or transmitter; the one at the operator's position is called the synchro motor or receiver. Other names frequently used for these devices are selsyn, autosyn, and synchrotie.

Synchros are used extensively in Air Force equipment. Almost all radar equipment having a remote indicator utilizes their principles. Other applications are remote readings in the radio compass and course direction in the omnidirectional receiver. It is also used in connection with the flux-gate compass, remote gun turrets, and radio-controlled missiles and aircraft.

A SIMPLE REMOTE-INDICATION SYSTEM

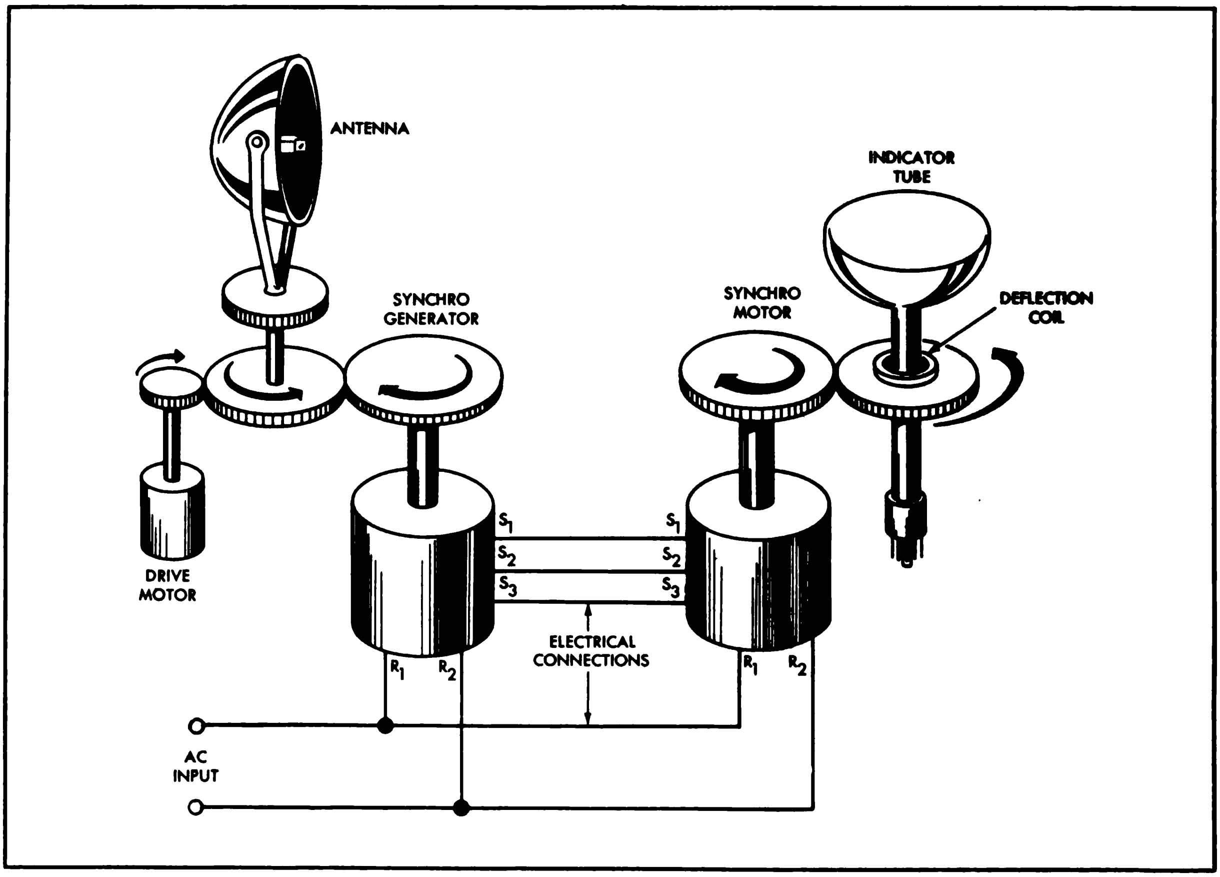
The synchronization of the turning of the deflection coils with the turning of the antenna in the radar set is accomplished by means of synchros. Two synchros comprise a simple system as shown in the illustration below. Note that one synchro is geared to the antenna motor and is driven by it. It is called the *synchro generator* or *transmitter* because it generates and transmits a rotational force (torque) to another synchro. The second synchro is called a *synchro motor* or *receiver* because it receives the rotational force and applies it to the deflection coil. The deflection coil then rotates in synchronization with the antenna.

Thus, a synchro may be described as an alternating-current machine which is used for the transmission of angular-position information. Another name commonly employed for a synchro is a selsyn. Selsyn is a contraction of the words self and synchronous and was originally used as a trade name by General Electric. However, the term selsyn is so widely used that it is now synonymous with synchro regardless of manufacturer. Another trade name for synchros (manufactured by Bendix) is autosyn. Synchros are also called synchronous or self-synchronous units. Regardless of name or manufacturer, the same general principles apply in all cases.

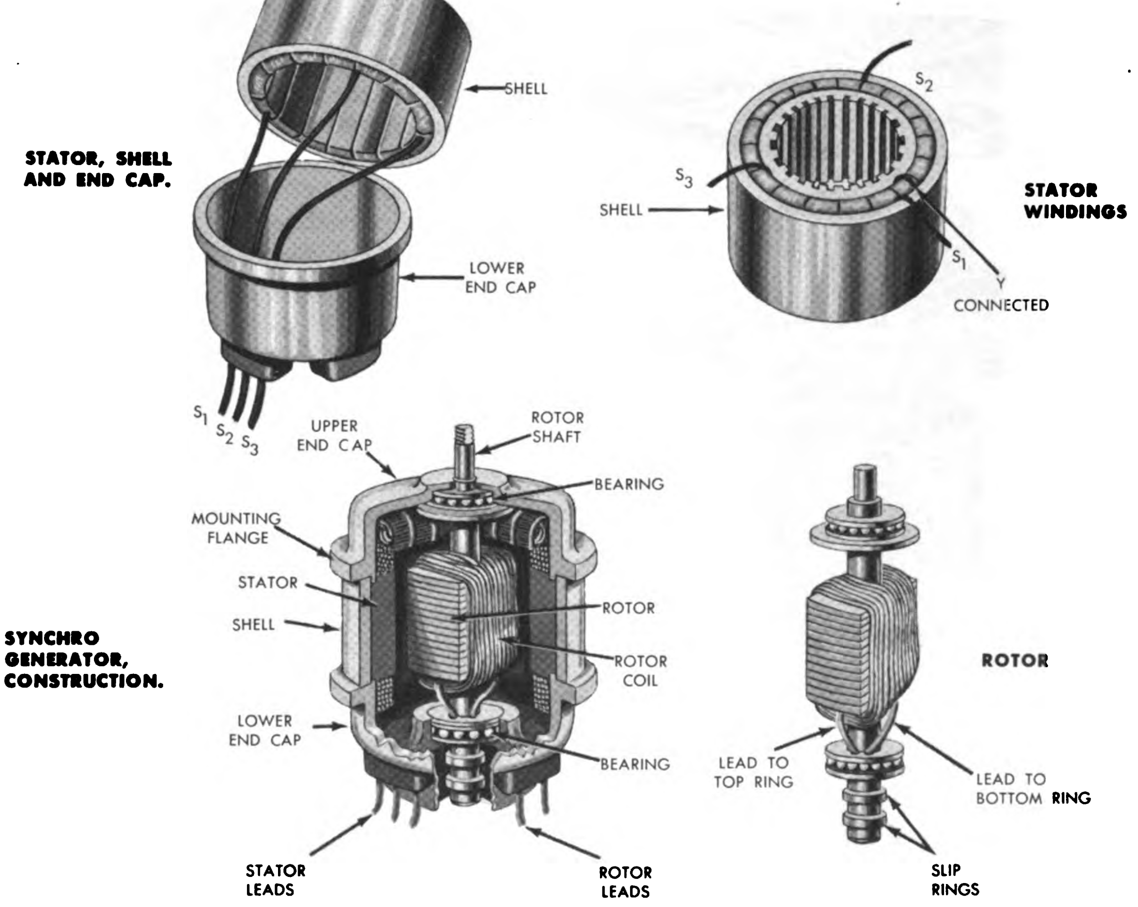
Synchros have a variety of uses in all types of electronic equipment. Wherever a remote indicating or control device is needed, an appropriate synchro may be found to do the job.

CONSTRUCTION OF SYNCHROS

Synchro generators and synchro motors are almost identical in construction. The synchro generator will be described and the differences between the motor and the generator pointed



Simple Synchro System

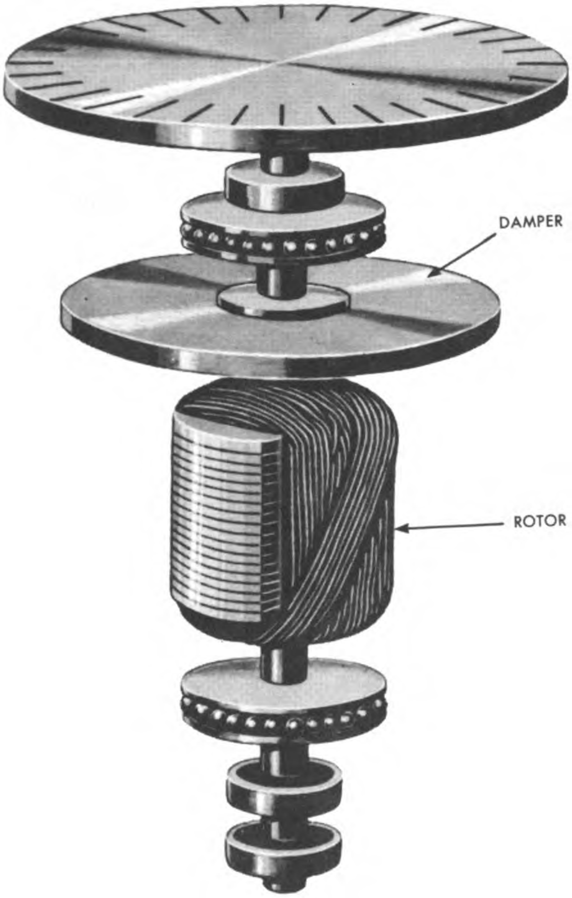


Construction of a Synchro Generator

out. The synchro generator resembles a small ordinary AC generator or motor. It has a stator and a rotor which are shown in the illustration above. The stator windings are composed of three groups of coils and are placed in slots around the inside of a laminated iron shell at 120° intervals. The three groups of coils are connected in a Y-connection like the stator of a three-phase Y-connected AC motor. However, *the voltages and currents present are single-phase, not three phase.* Note how the three stator leads are brought out through the lower end cap and labeled S_1, S_2, and S_3.

The rotor consists of two coils wound on a laminated soft-iron core and connected in series. The ends of the core are curved so that the air gap formed with the stator is small and uniform. The two leads from the rotor winding are connected to slip rings on the shaft and the entire rotor assembly is mounted on ball bearings to eliminate as much friction as possible. Note how the two external rotor leads (R_1 and R_2) are brought out through the lower end cap on the side opposite to the stator leads.

The stator of the synchro motor is identical to the stator of the generator; however, the rotor differs in that it has a heavy metal flywheel mounted on one end of the shaft. This flywheel, called an inertia damper, is shown in the illustration of the rotor of a synchro motor on the next page. It serves the purpose of preventing the motor from oscillating when the generator it is following stops or starts suddenly. The damper is mounted so that it can turn freely on the shaft for about 45° and then turns into a keyed bushing. This bushing is fastened to the shaft through a friction disk. For slow changes, the flywheel follows along; however, if a sudden



Synchro Motor Rotor

change is made in the position of the shaft, the flywheel tends to stand still, and the friction disk acts as a brake to stop the motion of the shaft.

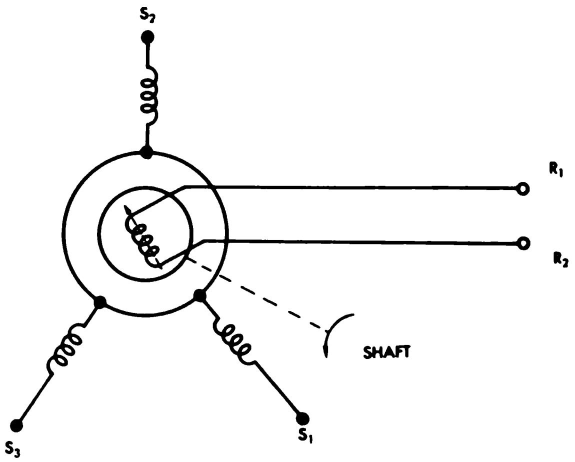
SYNCHRO DIAGRAMS

In order to make the operation of synchro more understandable, it is often desirable to employ schematic diagrams. The most common type of synchro schematic is shown in the illustration at the right. The stator coils are always shown with S_1, in the lower right position. Spaced at 120° intervals in a counterclockwise direction from S_1 are S_2 and S_3, respectively.

To make the description of the shaft positions simpler, one position is used as a reference and all others are measured in degrees of rotation from it. The reference position is called the electrical zero position and occurs when the rotor shaft is aligned with the axis of the S_2 winding. In measuring the number of degrees for any other position, you assume that you are looking at the shaft end of the rotor and that it is rotating counterclockwise from the zero position to 360°. The arrow shown at the R_1 end of the rotor points to the number which indicates the position of the rotor in degrees.

The rotor is always shown wound in the same direction from R_1 to R_2 regardless of its position. The stator coils are wound from the outside to the inside where they are connected to a common point. The open-headed arrows indicate the phase relationships of voltages. Arrows which point in the same direction indicate an *in-phase* relationship; arrows which point in opposite directions indicate a 180° out of phase relationship between voltages. Thus in the upper right diagram on page 477, the voltages from R_2 to R_1, from common to S_2, and from S_3 or S_1 to common are all in phase. The voltage from R_1 to R_2, however, is 180° out of phase with each of these voltages.

Although the schematic diagram discussed is most common, two other types of schematics are sometimes used for synchros. These are shown in the illustration on page 475. The lower schematic is especially useful in electro-mechanical circuits while the upper schematic is useful if you desire to show the details of the windings for explanation purposes. The schematic representations of a synchro generator and of a synchro motor are identical.



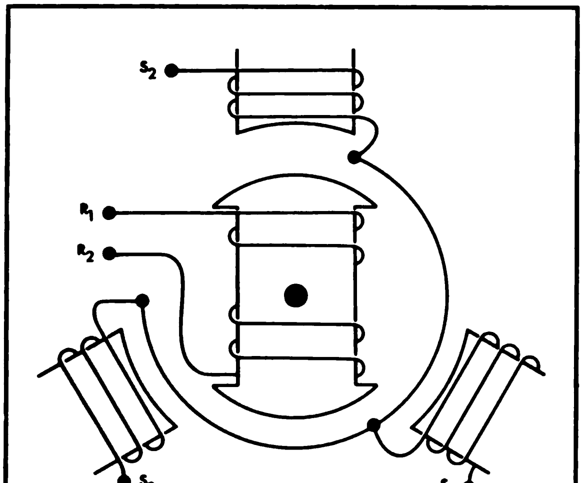
Wye-connected Stator

INDUCTION FROM ROTOR TO STATOR

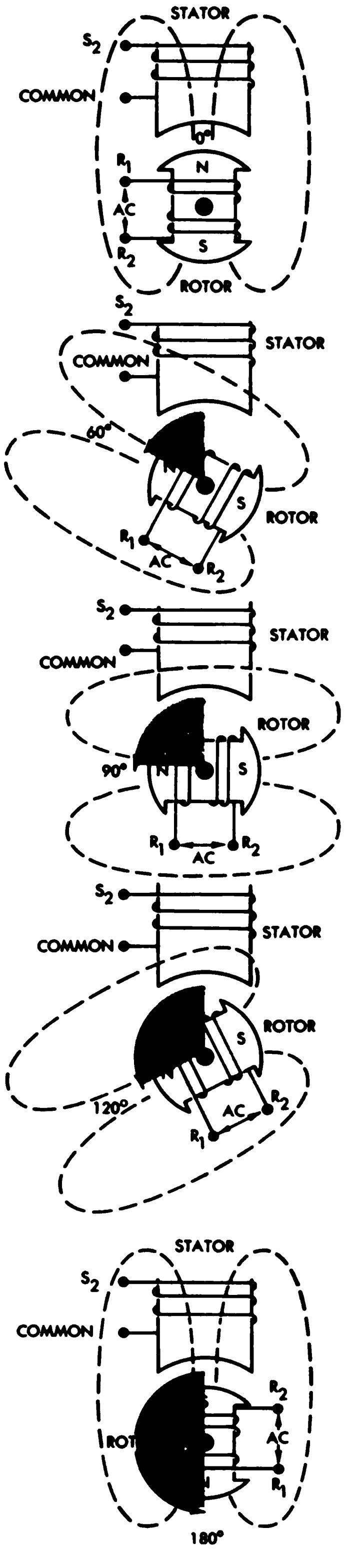
The operation of synchros is an application of transformer action with the rotor acting as a primary and each stator winding as a secondary. As the rotor turns to its different positions, it induces different voltages in each stator winding. Thus the voltages induced in the stator by the rotor indicate the position of the rotor at every instant. To understand how the position of the rotor changes the voltage induced in a stator coil, assume that the rotor is stopped in five different positions, as shown in the illustration to the right. In each case, consider that the polarity of the AC voltage applied to the rotor at the instant shown causes a magnetic flux to be set up with N and S poles as illustrated.

In the 0° position, all the flux lines (broken lines) of the rotor cut the windings of the S_2 coil. The voltage induced in this position is therefore a maximum and of such polarity as to oppose the N pole of the rotor. The voltage from R_1 to R_2 is in phase with this induced voltage from S_2 to common.

In the 60° position, only one-half the flux lines cut the windings of the S_2 coil. The voltage induced in this position is therefore one half the maximum voltage induced in the 0° position.



Synchro — Schematic Diagrams



Transformer Action in a Synchro

However, the polarity of the voltage is the same.

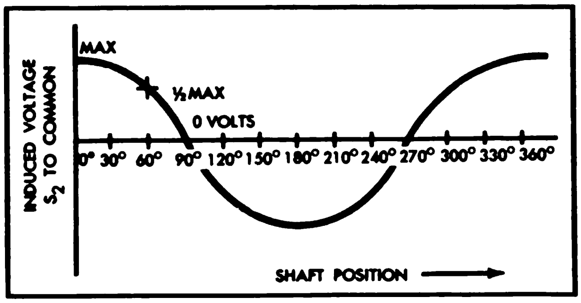
This is not a true explanation of the reason that no voltage is induced in the stator coil. Actually, the flux from the rotor induces equal and opposite voltages on the opposite sides of each turn. Thus, at this position, the net voltage is zero.

In the 120° position, only one-half the flux lines cut the S_2 stator windings. The voltage induced is therefore one-half maximum voltage just as in the 60° position. However, note that the S pole of the rotor is now closer to the S_2 coil and the voltage induced is of such polarity as to oppose the S pole of the rotor. This induced voltage is therefore 180° out of phase with the voltage induced in the 60° position and with the rotor voltage from R_1 to R_2.

In the 180° position, all the flux lines cut the S_2 stator windings. This induced voltage is 180° out of phase with the induced voltage in the 0° position and with the rotor voltage from R_1 to R_2.

As the rotor continues to turn, the induced voltage decreases from a maximum until it finally becomes zero at the 270° position. The induced voltage then begins to increase and is in phase with the voltage from R_1 to R_2 until at 360°, it is again a maximum.

Note that all voltages are either in phase or 180° out of phase with the rotor voltage from R_1 to R_2. If you plot induced voltage against shaft position, the curve shown in the illustration below results. Voltages shown above the axis are in phase with the rotor voltage from R_1 to R_2; those below the axis are 180° out of phase. From the curve, it is evident that the amount and phase of the induced voltage depend upon the position of the rotor. The curve resembles one made by plotting the values for the cosine of an angle from 0° to 360°. Therefore, the voltage induced in a stator is proportional to the cosine of an angle made by the rotor shaft and the axis of the stator winding. Expressed as a formula,



Induced Voltage in a Stator Winding

$$E_s = k \cos \theta$$

where E_s is the voltage induced in a stator winding, θ is the angle between the rotor shaft and the axis of the stator winding, and k is a constant which depends upon the specifications of the synchro.

NOTE

In aircraft installations where 115 volts AC is applied to the rotor, a common value for k is 52.

STATOR TERMINAL VOLTAGES

The general formula and curve for the voltage induced in a stator winding by a rotor were presented in the previous paragraph. If the angle θ is measured from the electrical zero position, however, the formula yields only the voltage induced in the S_2 winding. To obtain the values for the voltages induced in S_3 and S_1, subtract from θ 120° and 240°, respectively. Thus the formulas for the voltages induced in the three windings are:

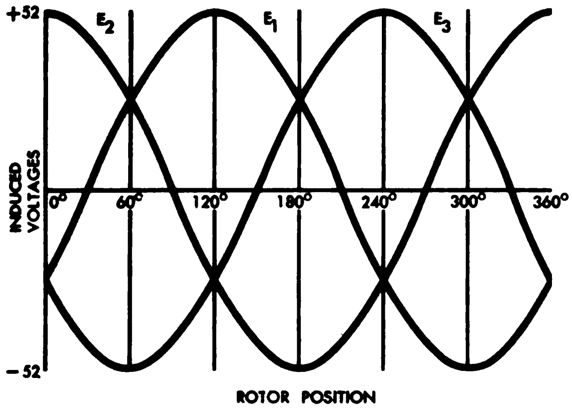
$$E_2 = k \cos \theta$$
$$E_3 = K \cos (\theta - 120°)$$
$$E_1 = k \cos (\theta - 240°)$$

where E_2, E_3, and E_1 are voltages induced in S_2, S_3, and S_1, respectively, and Θ is the angle of the rotor shaft measured from the electrical zero position.

Using $k = 52$ and plotting the values of E_2, E_3, and E_1 for values of θ in the appropriate formulas, you obtain the three curves shown in the illustration at the top of page 477.

Note that the voltage induced in each stator winding varies continuously from a maximum of 52 volts in one direction to zero to a maximum of 52 volts in the opposite direction.

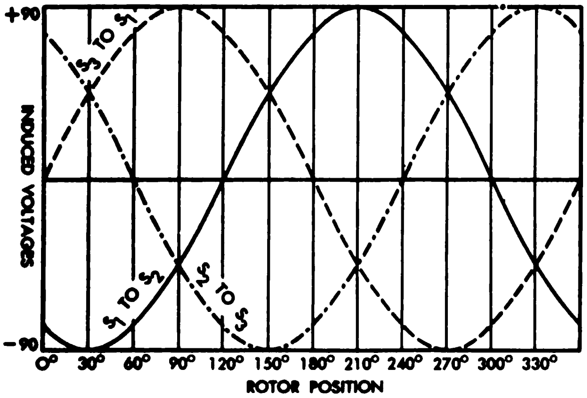
Since the terminals of a synchro are accessible for voltage measurements, it is of more practical value to know the voltages between terminals than that across each winding. In each case, the voltage between a pair of terminals is the vector sum of a pair of stator winding



Stator Voltages Induced by Rotor

voltages. The winding voltages to be added are either in phase or 180° out of phase; hence, the terminal voltages are either the sum or difference of a pair of winding voltages. For example, if you refer to the schematic diagram shown on page 475, you will see that the voltage between S_1 and S_2 is equal to the sum of the two stator winding voltages from S_1 to common and from common to S_2. Similarly, the voltage between S_2 and S_3 consists of the two stator winding voltages from S_2 to common and from common to S_3, and, the voltage from S_3 to S_1 consists of the two stator winding voltages from S_3 to common and from common to S_1.

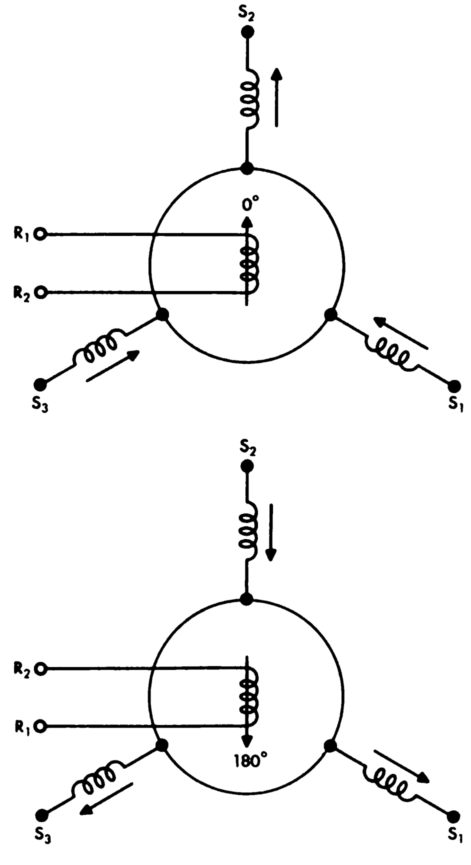
If, for various positions of the rotor in the illustration to the left, you add the appropriate pair of stator coil voltages, the voltage between terminals for each of these rotor positions is obtained. Three curves obtained by joining points found in this manner are shown in the illustration below. The curves indicate the voltages measured from S_1 to S_2, S_2 to S_3, and S_3 to S_1, respectively. Note that these voltages (for $k = 52$) vary continuously from a maximum of 90 volts in one direction to zero, to 90 volts in the other direction. As in the previous illustration, voltages in phase with the rotor voltage R_1 to R_2 are shown above the axis and those 180° out of phase are shown below the axis. Whether the synchro is a generator or a motor, the stator terminal voltages indicate the position of the rotor at any instant.



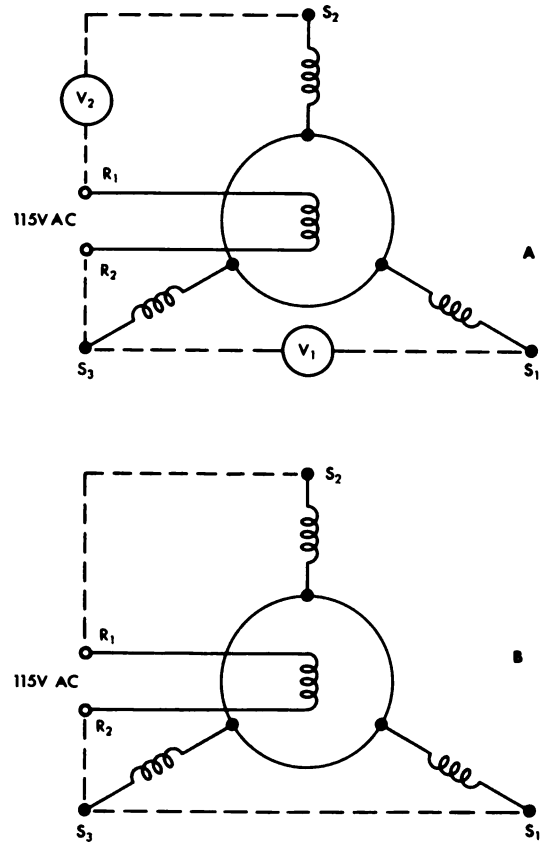
Terminal Voltages Induced by Rotor

ELECTRICAL ZERO

The measurement of the terminal voltages of a synchro is performed for various reasons, the most important of which is to find the electrical zero position of a rotor. The electrical zero position has already been defined as that position in which the rotor coil is lined up with the axis of the S_2 stator winding. In the illustration below, you can see that two such positions of the rotor are possible. In both the 0° position



Zero and 180° Rotor Positions



Finding Electrical Zero

and the 180° position, the terminal voltage from S_2 to S_3 (or S_1) is 78 volts and that between S_1 and S_3 is zero. However, in the 0° position the voltage from S_2 to S_3 (or S_1) is in phase with the voltage from R_1 to R_2 while in the 180° position the two voltages are 180° out of phase.

The electrical zero position is therefore completely defined as the position of the rotor in which the voltage between S_1 and S_3 is zero and the voltage from S_2 to S_3 (or S_1) is in phase with the voltage from R_1 to R_2.

Finding the Electrical Zero Position

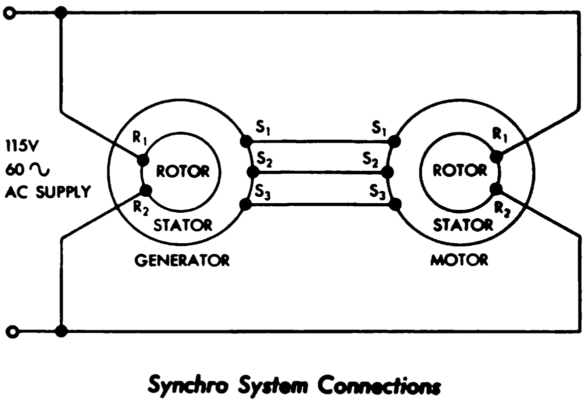
The electrical zero position of a synchro may be found in various ways. One common method is to connect AC voltmeters as shown at A in the illustration above. Turn the rotor until V_1 indicates zero. If V_2 indicates less than the applied voltage, the rotor is in the electrical zero position. If V_2 indicates more than the applied voltage, the rotor is 180° out of phase with the electrical zero position and should be turned 180°. For the synchro whose stator terminal voltages are shown by the curves on page 477, V_2 indicates 37 volts (115-78) in electrical zero position and 193 volts (115+78) in the 180° position.

If the meters are removed and the connections made as shown at B, the rotor assumes the electrical zero position when the AC voltage is applied to the rotor. However, you must be careful not to leave a synchro motor connected in this manner longer than is necessary to zero it or the motor will overheat and be damaged. This is due to the fact that 115 volts is being applied between S_2 and S_3 (or S_1) instead of the normal 78 volts.

OPERATION OF A SYNCHRO SYSTEM

When a generator and a motor are properly connected, you have a complete synchro system in which the motor shaft follows the generator shaft. The proper connections for a synchro system are shown in the illustration below. Note that the rotors are connected so that both R_1 terminals are joined to one side of the supply line and both R_2 terminals to the other. In this way, the two rotors have voltages applied which are identical in magnitude, phase, and frequency.

Each motor stator lead is joined to its corresponding generator stator lead, that is, S_1 to S_1, S_2 to S_2, and S_3 to S_3. Thus each stator winding voltage of the motor is opposed by each stator winding voltage of the generator. If both rotors are in corresponding positions, the induced voltages in the two S_1 coils are equal.



Synchro System Connections

The voltages induced in the two S_2 coils and in the two S_3 coils are likewise equal. Since these equal voltages oppose each other, the voltage between each pair of corresponding stator terminals is zero. Therefore, *no current flows in the stator windings when the two rotors are in corresponding positions.*

However, if the rotor of the generator is turned mechanically so that it no longer corresponds to the rotor position of the motor, the stator voltages of the generator are no longer equal to the stator voltages of the motor. Therefore, voltages exist between S_1 and S_1, S_2 and S_2, and S_3 and S_3. These voltages cause currents to flow in the stator windings, causing flux fields to be created which exert torque on the rotors. Because the generator rotor is held mechanically, the motor rotor is turned by this torque until the positions of the rotors correspond again and the stator voltages are equal.

Typical Stator Voltages and Synchro Action

Why the motor rotor follows the generator rotor may be understood more clearly if you assume two different conditions. First, assume that both rotors are in the electrical zero position as shown in the illustration below. The stator voltages are equal and oppose each other. Therefore, no current flows in any of the stator windings and no torque is produced. The rotors remain stationary.

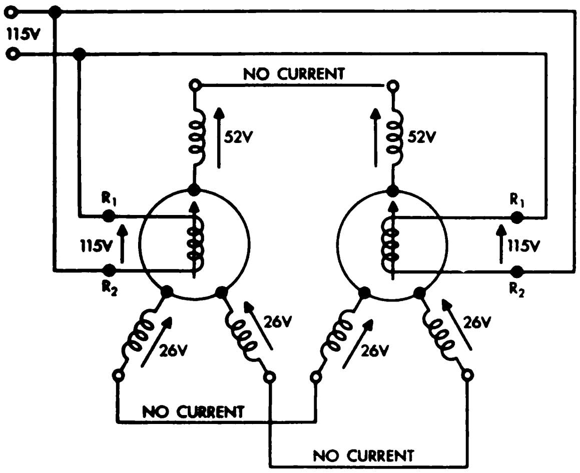
Second, assume that the generator rotor is turned 30° from electrical zero. This causes the stator voltages of the generator to change as shown in the illustration below. Note that a balanced condition no longer exists and current flows in all three stator leads and coils. The amount of current flow in each stator lead is shown in the diagram. At one instant, current in each stator lead flows as follows:

In the S_1 leads—From common in the motor down through the S_1 coil, up through the S_1 coil of the generator to common.

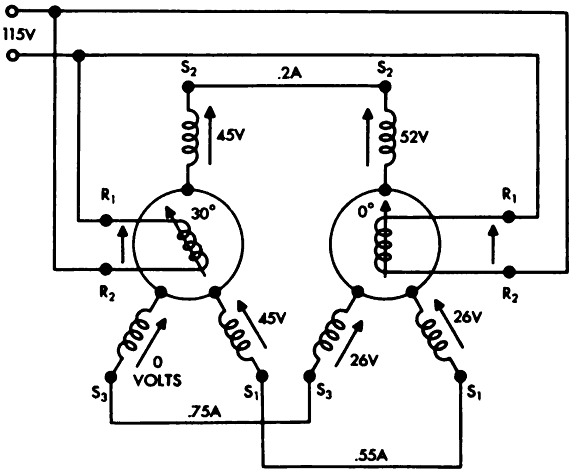
In the S_2 leads—From common in the motor up through the S_2 coil, down through the S_2 coil of the generator to common.

In the S_3 leads—From common in the generator down through the S_3 coil, up through the S_3 coil of the motor to common.

The currents through the stator coils create magnetic fields whose poles are as shown in the illustration on the next page. Check these magnetic poles by applying the left-hand rule using the direction of the current shown by the arrows. Note that the magnetic poles of the two rotors are identical. However, the magnetic poles of the generator stator coils are opposite to those of the motor. The torque exerted is therefore in a clockwise direction for the generator and in a counterclockwise direction for the motor. Either movement would tend to align the two rotors; however, the generator is mechanically driven and is not free to rotate. *The rotor of the motor therefore turns in a counterclockwise direction and aligns itself with the rotor of the generator.*



Generator and Motor in 0° Position



Shafts in Different Positions

NOTE

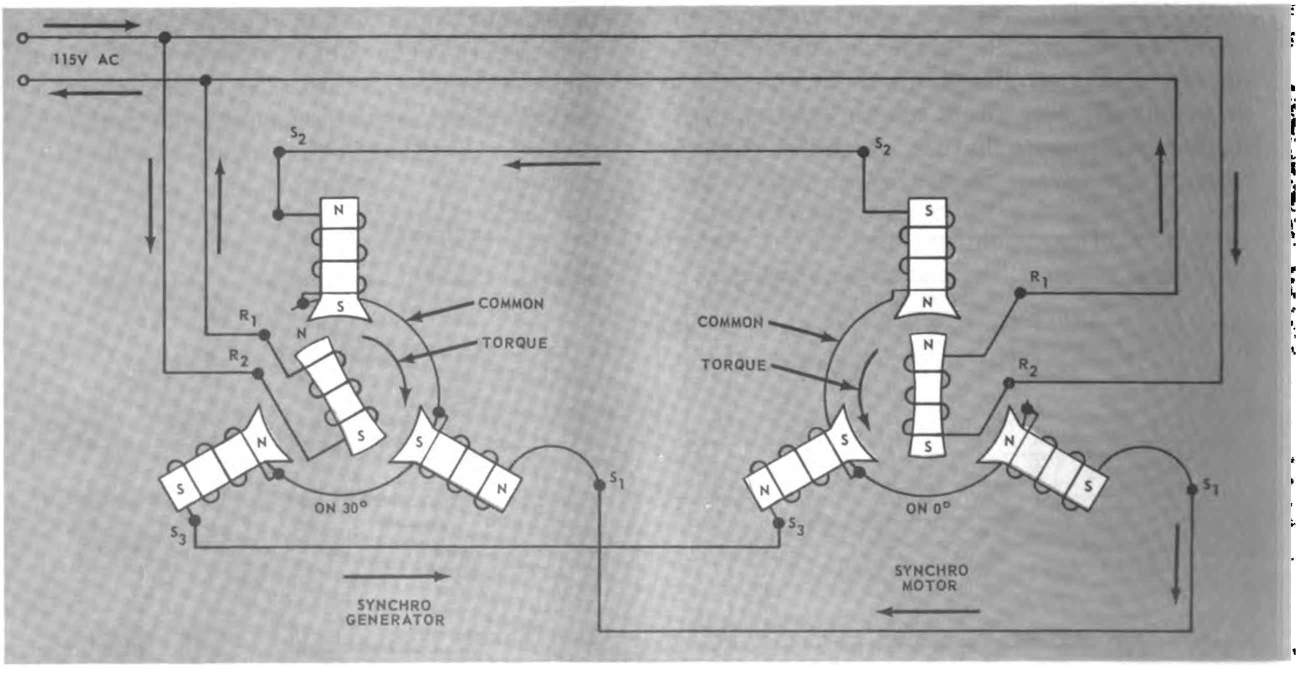
The direction of the torque developed remains the same if the input voltage is reversed in phase. This reversal in phase causes the directions of all the currents, and therefore all the magnetic poles, to reverse at the same time. The direction of the torque, however, remains the same.

Variation of Torque with Angular Displacement

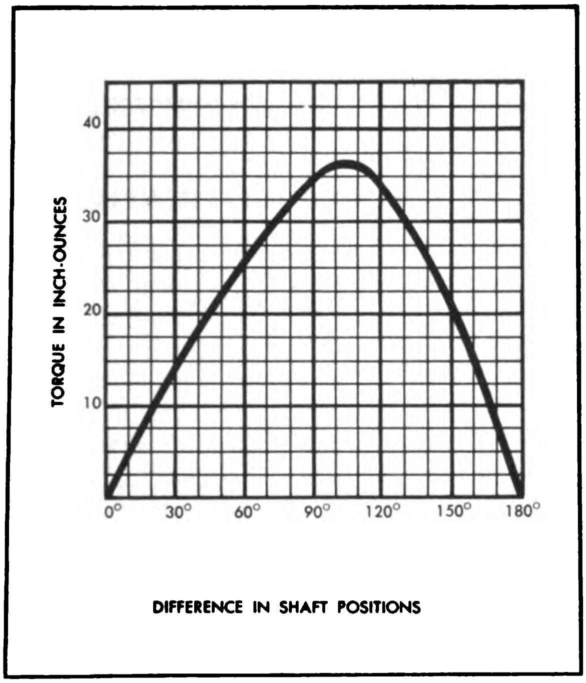
The voltages induced in the stator coils when the positions of the generator rotor and the motor rotor correspond are equal and of such polarity as to oppose each other. No current flows in the stator windings for this reason. However, if the generator rotor is displaced, the balanced condition is upset and currents flow in the two stators causing two opposing torques to be exerted. Since the generator rotor is the driven element, the motor turns to correct the displacement. As the motor turns and the two shafts approach correspondence, the stator currents become less and less. Finally, at correspondence, the stator currents are zero. The torque exerted on the rotor is nearly proportional to the stator currents for small angles; therefore, the torque also becomes less and less as the rotors approach correspondence.

The relationship between torque and the angular displacement between rotors is shown in the illustration at the top of the next page. It is evident that maximum torque occurs at a displacement angle of approximately 100° and that the torque is very weak at angles close to correspondence. Because of the weak torque near correspondence, synchro motors are used only for positioning light loads, such as a pointer or a dial. Loads are normally made so small that the displacement angle does not exceed an angle of 10° or 20°.

Even under no-load conditions, friction is always present in the motor in spite of the fact that ball bearings are finely machined and selected with great care. To overcome this drag caused by friction, torque must be produced. Since torque is produced only when the generator and motor are not in correspondence the motor must always lag slightly behind the generator. The accuracy of a synchro system can therefore never be perfect. However, by



Instantaneous Magnetic Fluxes Created by Currents in Stator



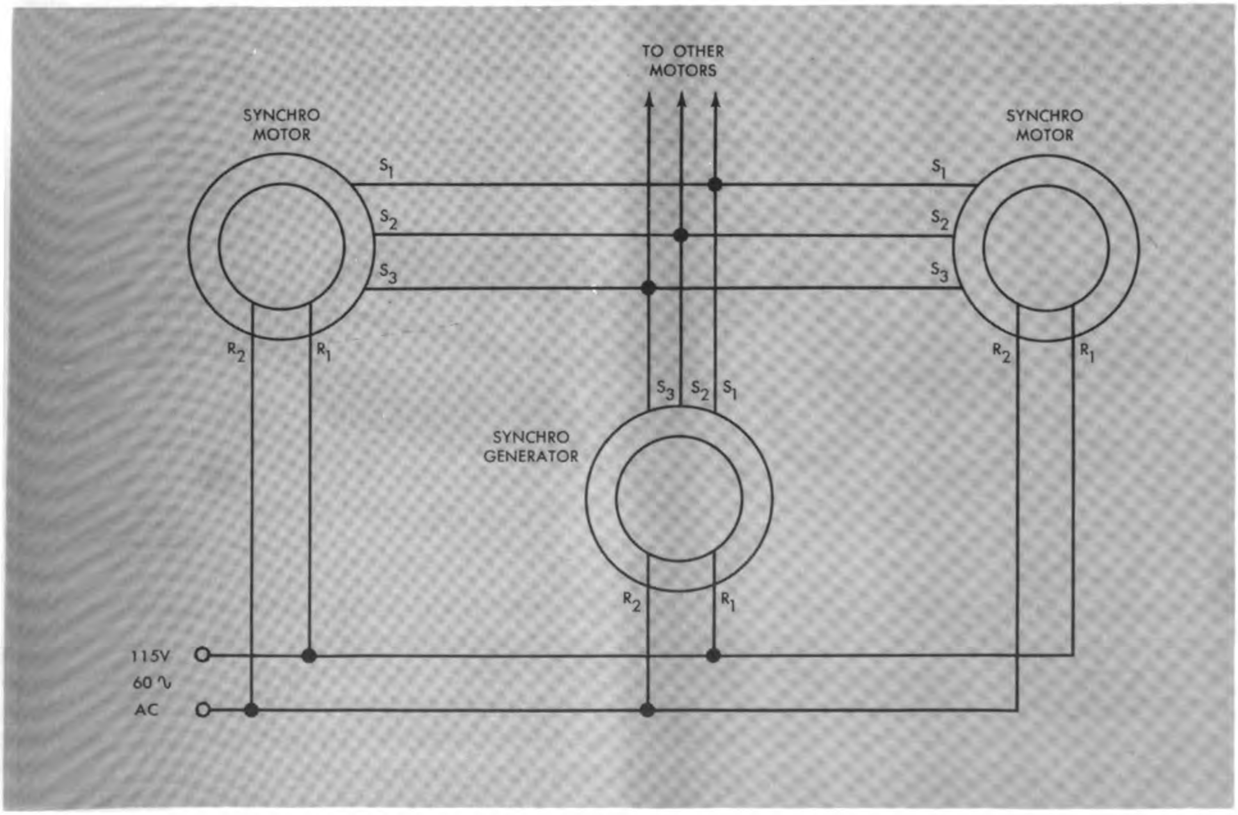
Torque and Displacement Angle

reducing friction to a minimum and producing sufficient torque to overcome this friction within a small angle of lag, synchros achieve a great degree of accuracy. A motor accuracy within a half degree or better is common in synchro systems.

System with Several Motors

One synchro generator is sometimes used to transmit positioning data to several motors located at different places. The standard type of connection is shown in the illustration below. Note that all the R_1 leads of the rotors are connected to one side of the input line, and all the R_2 leads to the other. The stator leads are connected in parallel across the generator stator leads.

In such a system, a larger generator is required to maintain a high degree of accuracy. If too many motors are connected, output voltages of the generator are reduced and excessive lag occurs in all the motors. If any one motor turns hard or becomes jammed, the accuracy of the whole system is affected.



Synchro Connections for Several Motors

Effect on Synchro System of Improper Connections

A synchro system is affected if the leads are not connected in the proper manner. These connections may have been made purposely to secure a desired effect, or accidentally. In the latter case, the recognition of the resulting symptom may lead the mechanics to the discovery of the incorrect connections. Three of the most common connection changes and the results they produce are:

The rotor connections reversed so that R_1 *is connected to* R_2. This results in a shift of 180° in the rotor of the motor. The motor then follows the generator but remains 180° displaced from it.

A cyclic shift of stator connections: that is, S_1 *to* S_2, S_2 *to* S_3, *and* S_3 *to* S_1. This results in a 120° counterclockwise shift of the motor. The motor then rotates in synchronization with the generator but remains 120° from it. A shift of two stator terminals so that S_1 is connected to S_3, S_2 to S_1, and S_3 to S_2 produces a similar effect but the angle between generator and motor positions is 240°.

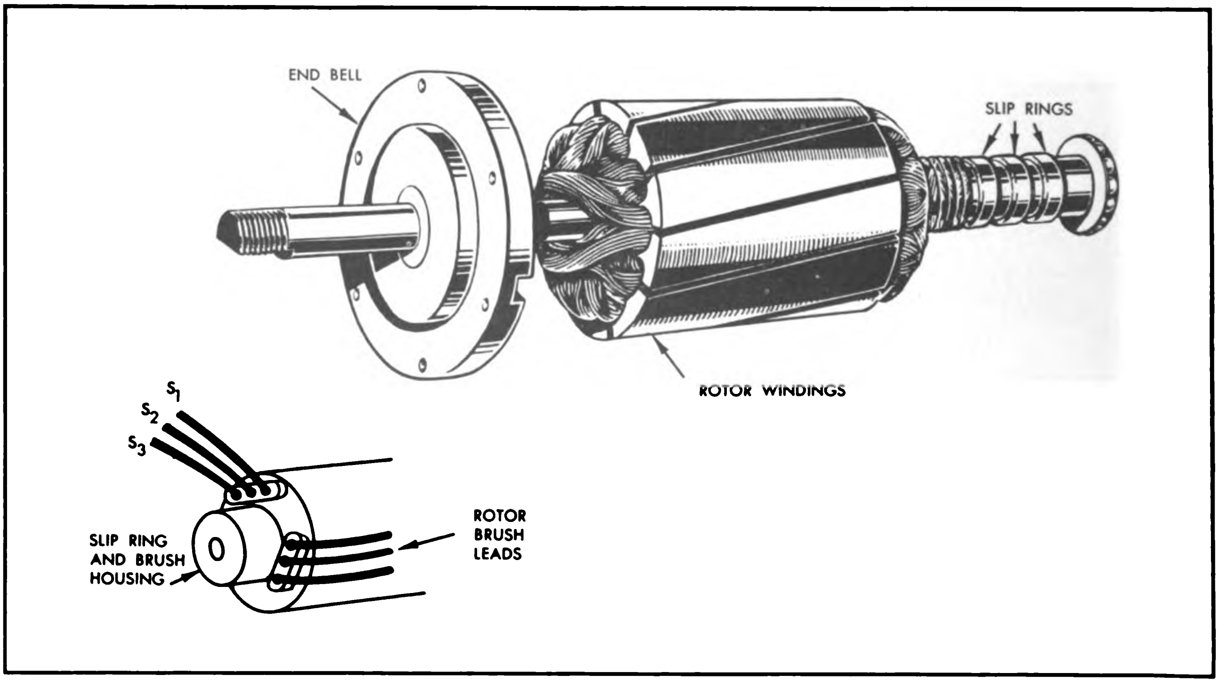
Interchanging two stator leads. This causes the motor to rotate in a direction which is opposite to that of the generator. When this is done purposely, the leads interchanged are S_1 and S_3, and the position of the motor at any instant is obtained by subtracting the generator position from 360°.

THE SYNCHRO DIFFERENTIAL

Synchro differentials, or differential selsyns as they are sometimes called, perform the same function electrically that a mechanical differential performs in a mechanical system. The differential in a mechanical system connects three shafts so that one shaft turns an amount which is equal to the difference between the amounts that the other two turn. A synchro differential generator subtracts two inputs and a synchro differential motor indicates the difference between two inputs. The action of a synchro differential may be reversed so that two inputs are added by a generator or the sum of two inputs indicated by a motor.

Construction of a Synchro Differential

As in the ordinary synchros, the differential generator and differential motor are alike electrically; however, the rotor of the motor has an inverted damper. This damper serves the



Rotor of Synchro Differential

same purpose as it does in the ordinary synchro; that is, it damps overdrive and prevents oscillation.

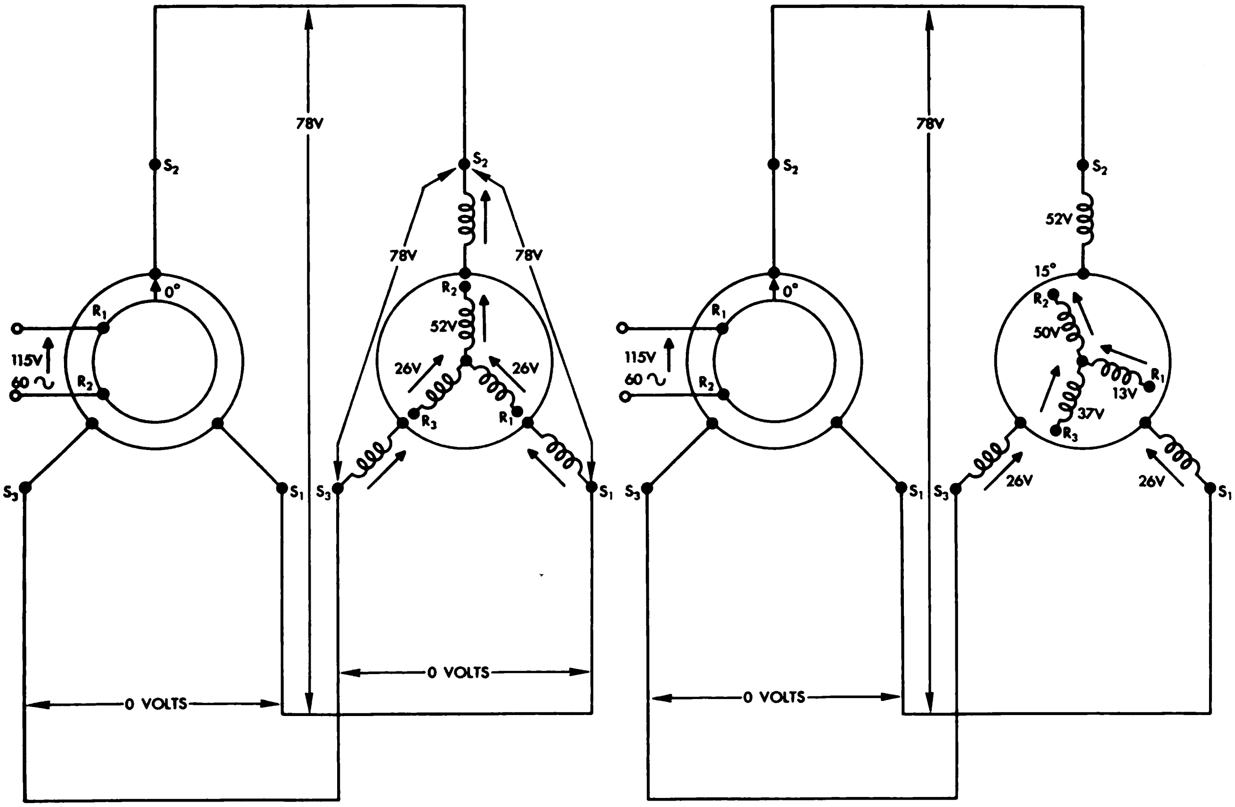
The stator of a differential generator or motor is similar to that of an ordinary synchro. It consists of three sets of Y-connected coils wound in slots which are spaced 120° apart around the inside of the field structure. The rotor of the differential, on the other hand, is completely different from that of the ordinary synchro. It is cylindrical in shape and has three sets of Y-connected coils wound in slots which are spaced 120° apart around the circumference. As you can see in the illustration shown on the preceding page, three slip rings connect the rotor leads R_1, R_2, and R_3 to the external circuit.

An external view of the brush end of the differential is shown in the inset. There is no AC voltage applied to the rotor leads as in the ordinary synchro. In the electrical zero position, the rotor coils are turned so that R_1, R_2, and R_3 are aligned with the stator coils S_1, S_2, and S_3, respectively.

Transformer Action

Like the ordinary synchros, the differential operates on the transformer principle. The stator acts as the primary and the rotor as the secondary of a 1:1 transformer. Because of the air gap, more turns are wound on the stator coils than on the rotor coils in order to achieve the 1:1 ratio. Therefore the stator must always be used as the primary and the rotor as the secondary of a differential.

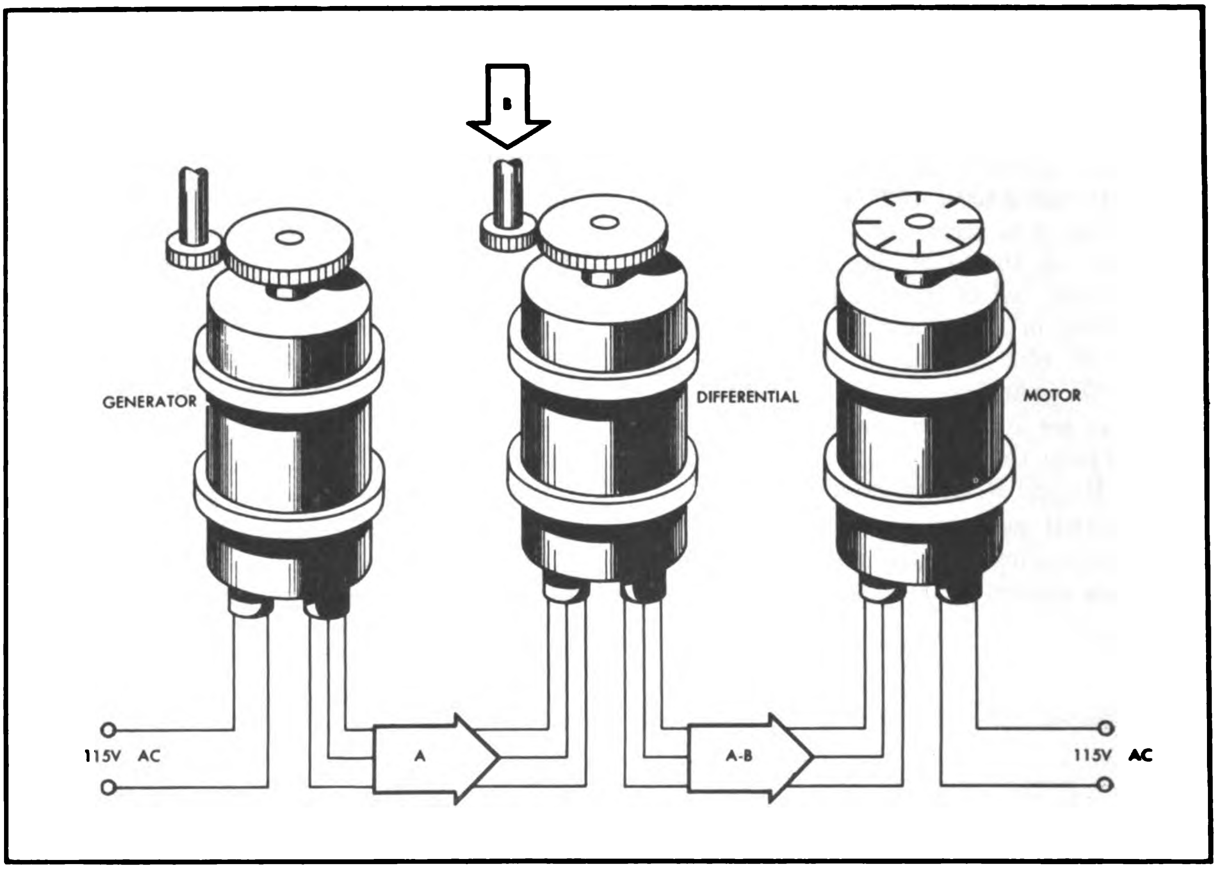
The transformer action of the synchro differential is illustrated in the diagrams below. Assume that both rotors are in the electrical zero position, the stators are connected properly as



0° POSITION

15° POSITION

Transformer Action in a Differential



Action of a Differential Generator

shown, and the rotor of the differential is open. Since the stators are connected together and are in parallel, the voltages in both sets of stators are equal to each other and are in phase. The stator voltages of the differential, through transformer action, cause similar voltages to be induced in each of the rotor windings. If these rotor windings were connected to the stator windings of a synchro motor and AC applied to its rotor, this motor would align itself in the electrical zero position. Note that this action is the same as though the differential were not in the circuit.

If the rotor of the differential were turned to the 15° position and the rotor of the synchro generator held in the electrical zero position as shown, the stator voltages are correspondingly equal as before. However, the voltages induced in the rotor coils are different because all the flux from each stator winding no longer cuts its corresponding rotor windings. These induced voltages as shown on the rotor windings are: 63 volts between R_1-R_2, 87 volts between R_2-R_3, and 24 volts between R_3-R_1. If the rotor windings were connected to the stator windings of a synchro motor, the synchro rotor would turn to the 345° position. This may be verified from the curves of stator voltages vs. rotor position shown on page 477. Note that a 15° counterclockwise rotation of the rotor of the differential generator is equivalent to a 345° counterclockwise, or 15° clockwise, rotation of the rotor of the synchro motor.

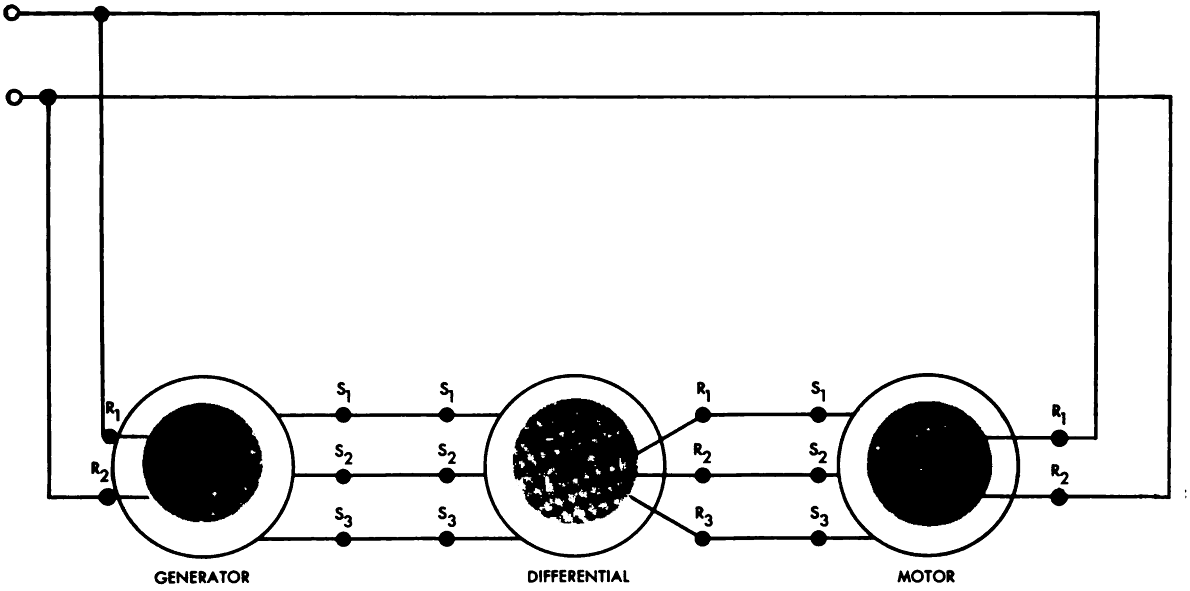
Action of a Differential Generator

A synchro differential generator receives two inputs; one is electrical and the other is mechanical. It subtracts one input from the other and transmits the difference. The action of a differential generator is shown in the diagram above. The electrical signal (A) from the stator or the synchro generator is applied to the stator of the differential. The mechanical signal (B) is obtained from a shaft which is connected to and turns the rotor of the differential. The differential subtracts the mechanical signal B from the electrical signal A and transmits the resultant signal to the stator of the synchro motor. The rotor of the motor then turns to a position such that it indicates A minus B.

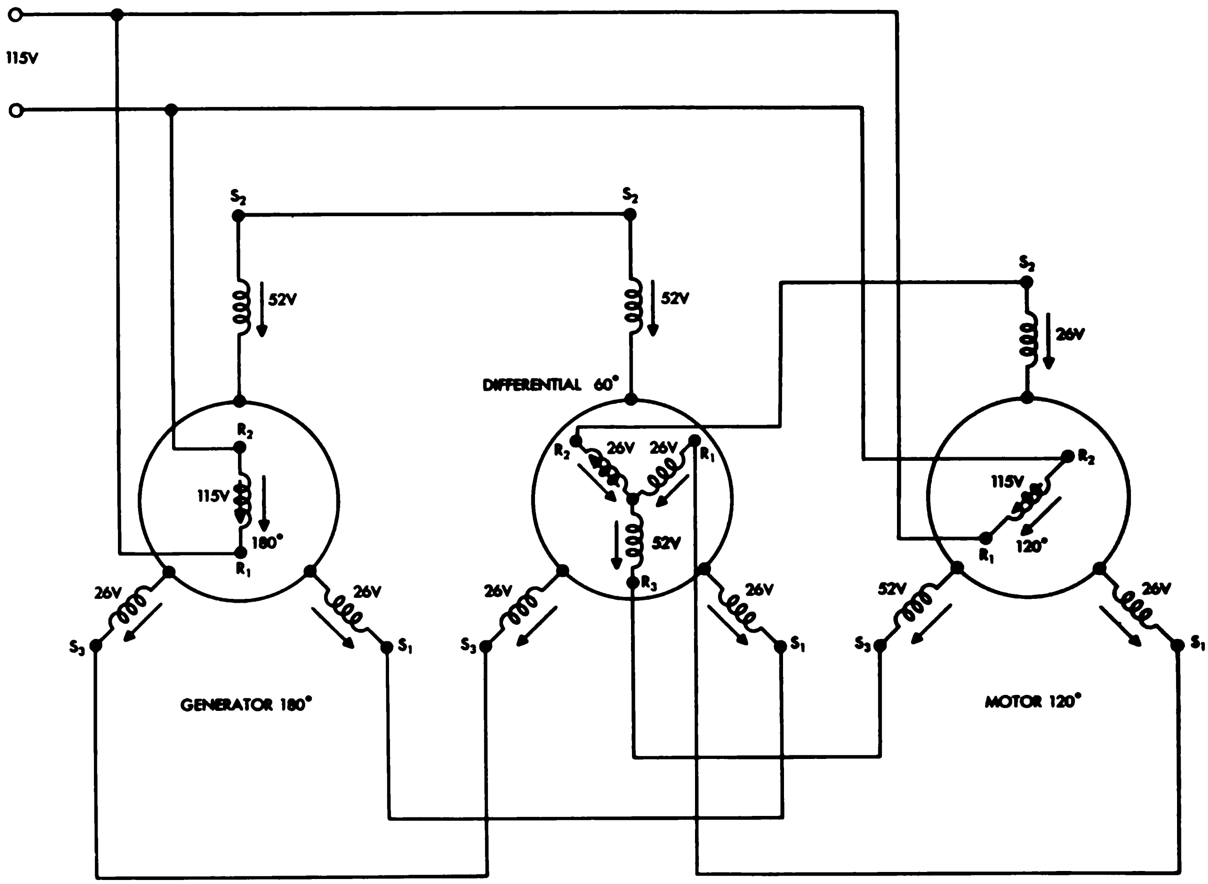
In order to indicate a difference position, a differential generator must be connected as shown schematically in the illustration below. Note that the differential generator stator leads are connected to the corresponding synchro generator stator leads and that the differential generator rotor leads are connected to the corresponding synchro motor stator leads.

In the preceding example where the differential generator was turned to a 15° position, note that the synchro motor turned to the 345° position. This is a difference reading since 15° subtracted from 0° (or 360°) is equal to −15° (or 345°). The difference reading is obtained whether both rotors are turned or only one is turned. The illustration on the next page shows a differential synchro system in which the rotor of the synchro generator is turned 180°, and the rotor of the differential generator is turned 60°. The resulting positioning of the rotor of the synchro motor at 120° (180° − 60°) is achieved as follows:

The rotor of the synchro generator, which is at 180°, induces voltages in the stator windings of the synchro generator. The amount and phase of these voltages are shown on the schematic diagram Differential Synchro System. Identical stator voltages are measured across the stator windings of the differential generator. The resultant magnetic field created by these stator voltages is equivalent to an axis of 180°. If the rotor of the differential generator were in its 0° position, identical voltages would be induced in the rotor windings and result in a 180° positioning of the motor. However, the rotor of the differential has been shifted 60° counterclockwise. This is equivalent to a shift of 120° from the magnetic field created by the stator. Therefore, the voltages induced in R_1, R_2, and R_3 are 52 cos 120°, 52 cos 240°, and 52 cos 0°, respectively. As shown in the diagram, the voltages between R_1-R_2, R_2-R_3, and R_3-R_1 are 0 v., 78 v., and 78 v., respectively, with phase relationships as represented by the arrows. When these voltages are placed across the stator windings of the motor, the rotor aligns itself in the 120° position. This may be



Differential Generator Connections



Differential Synchro System

verified by checking the curves of terminal voltages vs. rotor position shown in illustration on page 477.

Reversed Differential Action

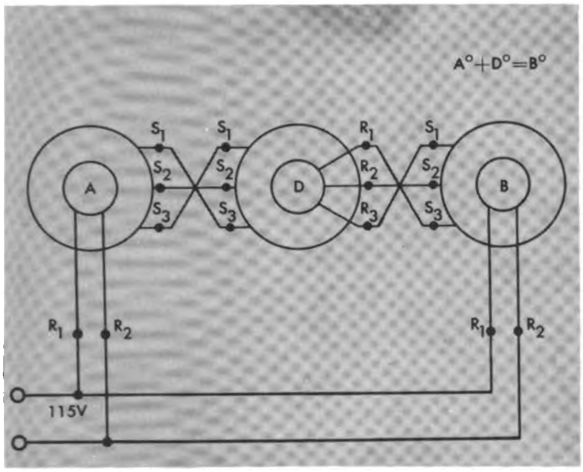
It is possible to reverse the action of a differential so that it will add two inputs. As before, one input is electrical and the other mechanical. The method of connecting a differential for addition is to reverse the S_1 and S_3 stator leads as well as the R_1 and R_2 rotor leads. The illustration on the next page shows a differential connected so as to add inputs. The differential that is connected in this manner will add two inputs whether both or only one of the rotors is turned.

The Synchro Differential Motor

The synchro differential motor indicates the difference between two inputs. The inputs are received from two synchro generators; one is applied to the stator and the other to the rotor of the differential motor. The motor subtracts one input from the other and positions its shaft to indicate the difference on a dial. A differential motor will indicate the sum of two inputs if its connections are reversed, that is, S_1 for S_3 and R_1 for R_3.

CONTROL TRANSFORMER

A control transformer is similar to a synchro motor in that it receives a signal from either a synchro generator or a differential generator. However, it differs in that its output is a voltage and not a shaft position. Actually, it is a transformer whose secondary may be moved. It transforms an electrical input into an electrical output that may be used to control the positioning of some device so that it will correspond with that of an input shaft.



Differential Reversed

Construction of a Control Transformer

The stator of a control transformer is similar to that of other synchros in that it has three windings that are Y-connected and spaced 120° apart. However, the windings differ in that there are more turns and they are constructed of finer wire than those of the ordinary synchros. The impedance of the stator of the control transformer is therefore higher and offers more opposition to the flow of current.

The rotor of the control transformer resembles the rotor of the synchro differential in physical appearance as shown below. However, the windings are connected in series and have only two external connections through slip rings like the rotor of the ordinary synchro. Very little current flows in the rotor windings because of the high impedance offered by the many turns of fine wire of which it is constructed.

The stator windings act as the primary and the rotor windings as the secondary of a 3:2 transformer.

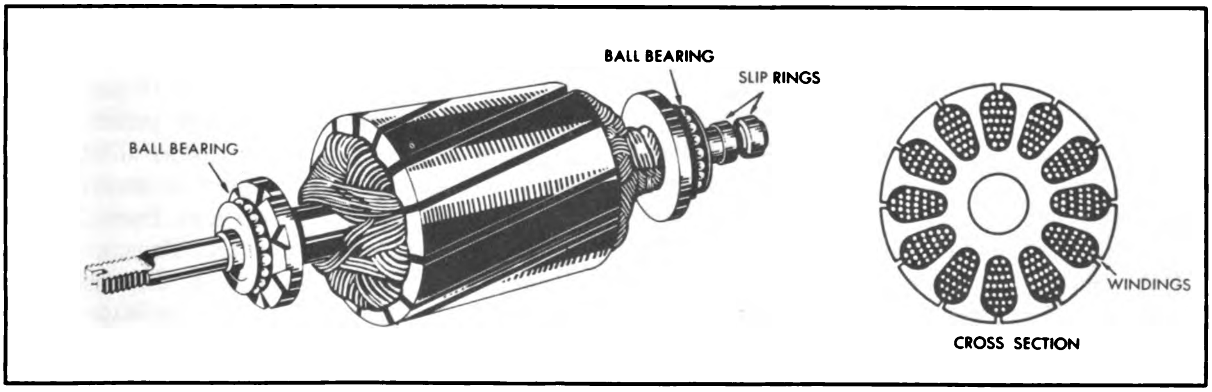
Schematics and Electrical Zero

Two types of schematic representation for control transformers are shown in the illustration on the next page. The upper is the more usual type which is found in most drawings; the other shows the rotor and stator windings in more detail.

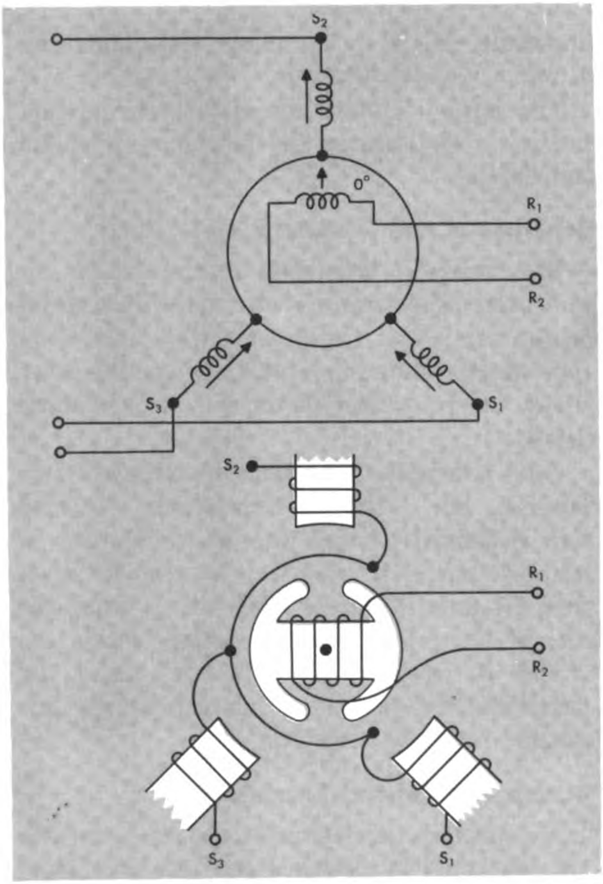
Both schematics represent the electrical zero position. For a control transformer, electrical zero is defined as that *position of the rotor in which no voltage is induced in the rotor windings from the stator windings.* This occurs when the axis of the rotor is perpendicular to the axis of the S_2 winding. Note that terminals are designated in the same manner as for the synchro motor or generator.

Action of a Control Transformer

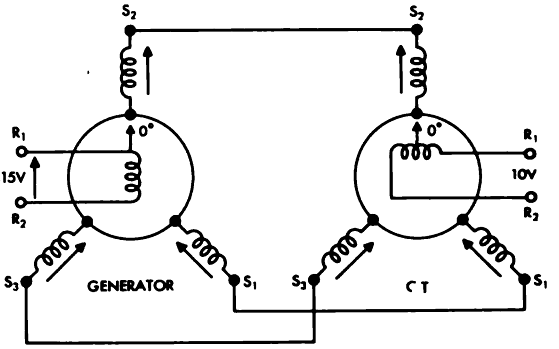
A control transformer and a synchro generator are connected in electrical zero position. Note the relative positions of the two rotors as shown in the illustration on page 488. No voltage is induced in the rotor of the control transformer by the S_2 coil since the rotor and coil are mutually perpendicular. The currents in the S_1 and S_3 coils are equal and opposite and therefore induce voltages in the rotor which are equal and opposite. The net effect is that *no voltage is induced in the rotor in this position.*



Typical Control Transformer Rotor



CT Schematics



Electrical Zero Position

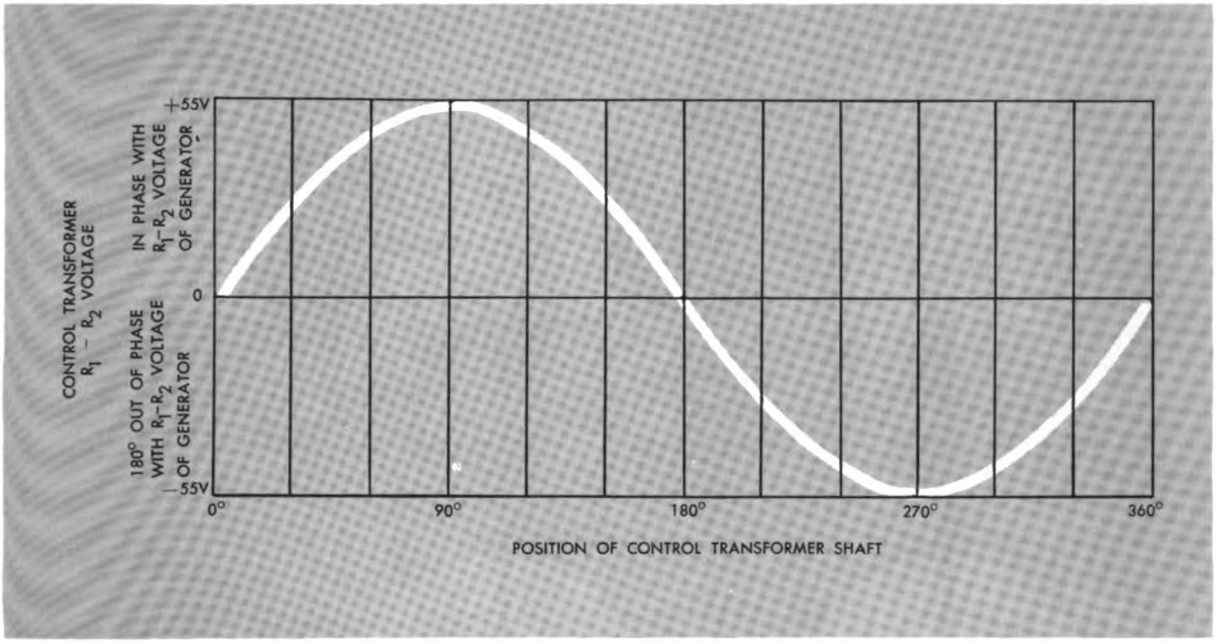
Turn the shaft of the control transformer to the 90° position. Since the S_2 coil is aligned with the rotor, it induces maximum voltage in the rotor. The currents in the S_1 and S_3 are equal and in the same direction. They therefore induce voltages in the rotor which are equal and of the same polarity. The net effect is that *maximum voltage is induced in the rotor in this position* and this induced voltage is *in phase* with the voltage of the rotor of the generator.

Turn the shaft of the control transformer to the 180° position. The S_2 coil and the rotor are again mutually perpendicular and the currents in the S_1 and S_3 coils are again equal and opposite although they are reversed. The net effect is that *no voltage is induced in the rotor in this position.*

Turn the shaft of the control transformer to the 270° position. The S_2 coil is again aligned with the rotor and the currents in S_1 and S_3 are again equal and in the same direction. The net effect is that *maximum voltage is induced in this position;* however, this induced voltage is 180° *out of phase* with the voltage of the rotor of the generator.

In all intermediate positions of the shaft, the magnitude of the voltage induced in the rotor varies with the position of the shaft. The induced voltage is in phase with the voltage of the rotor of the generator for shaft positions from 0° to 180° and 180° out of phase for shaft positions from 180° to 360°. The variation of induced voltage with position of control transformer shaft is shown in the illustration on the next page. With the rotor of the generator held in zero position, this variation is a sine wave with voltages that are in phase with the rotor of the generator shown above the axis and voltages 180° out of phase shown below the axis.

If the rotor of the synchro generator is turned from the 0° position, the voltage induced in the rotor of the control transformer becomes zero when it is moved to a corresponding position. With this new position used as a reference, the voltages induced in the rotor of the control transformer may also be represented by a sine wave similar to the one for the 0° position. In this case, the induced voltages are in phase with the voltage of the rotor of the generator for positions of the control transformer shaft which are 0° to 180° from the reference position and 180° out of phase for positions from 180° to 360° from the reference position. Thus, a *counterclockwise rotation* (0° to 180°) of the shaft from the reference position yields a voltage across the rotor of the control transformer which is *in phase* with the voltage of the generator rotor, and a *clockwise rotation* (0° to 180°) of the shaft yields a voltage which is 180° *out of phase.*



Induced Voltage versus Shaft Position

NOTE

The output voltage of a control transformer is zero when its rotor position corresponds, or is 180° out of phase, with the rotor position of the generator. To verify the fact that the two rotor positions correspond when the output voltage is zero, turn the shaft of the control transformer slightly counterclockwise. If the output voltage is in phase with R_1-R_2 of the generator, the two rotor positions correspond; if 180° out of phase, the two rotor positions are 180° out of phase.

The following important facts summarize the action of a control transformer:

If the shaft of a synchro generator and that of a control transformer correspond, there is no voltage output.

If the synchro generator shaft is turned clockwise (equivalent to ccw rotation of the control transformer, an output voltage is obtained which is in phase with R_1-R_2 of the generator and whose magnitude depends upon how much the two shafts are out of correspondence.

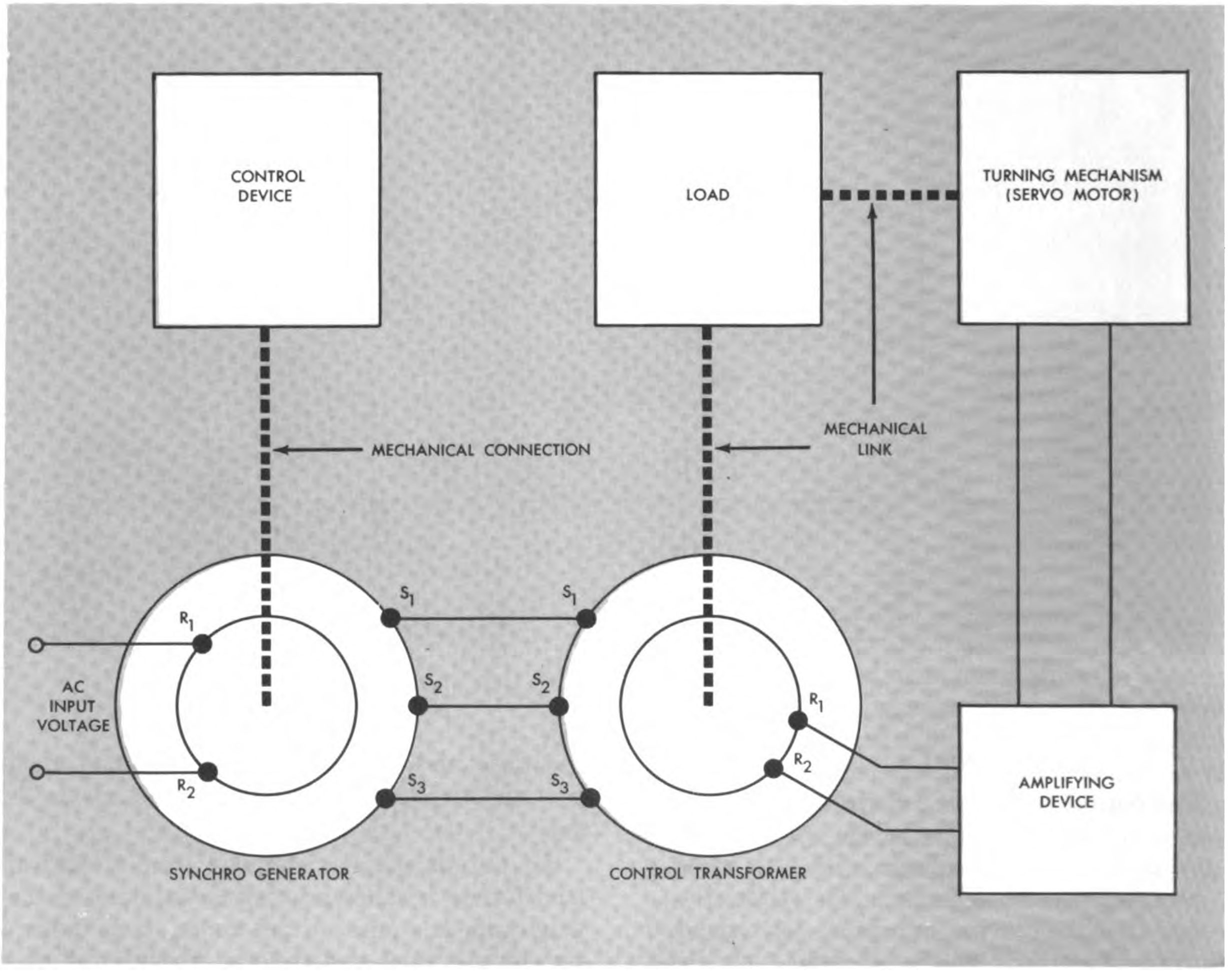
If the synchro generator shaft is turned counterclockwise, an output voltage is obtained which is 180° out of phase with R_1-R_2 of the generator and whose magnitude depends upon how much the two shafts are out of correspondence.

An output voltage obtained from a control transformer is called an *error signal* because the magnitude and phase of this voltage is an indication of *how much* and *in what direction* the two rotors are out of correspondence. Turning the rotor of the control transformer so that its position corresponds with that of the generator rotor cancels or *nulls* the error signal. The position of correspondence is sometimes referred to as the *null* position.

How the error voltage may be applied to turn a device so that its position corresponds with the position of the rotor of a synchro differential is explained in the discussion of servomechanisms.

SERVOMECHANISMS

A synchro system is a system for automatically controlling the rotation of a shaft so that it will turn in correspondence with another shaft which is remote from it. The shaft being turned in such a system is usually attached to a very light load such as an indicating device or dial. Therefore very little torque is required. A *servomechanism* or *servo* is also a system for automatically con-



A Simple Servomechanism

trolling the rotation of a shaft so that it will turn in correspondence with another shaft which is remote from it. However, the shaft being turned is attached to a greater load making it necessary for a *torque-amplifying* device to be in the system.

A Simple Servomechanism

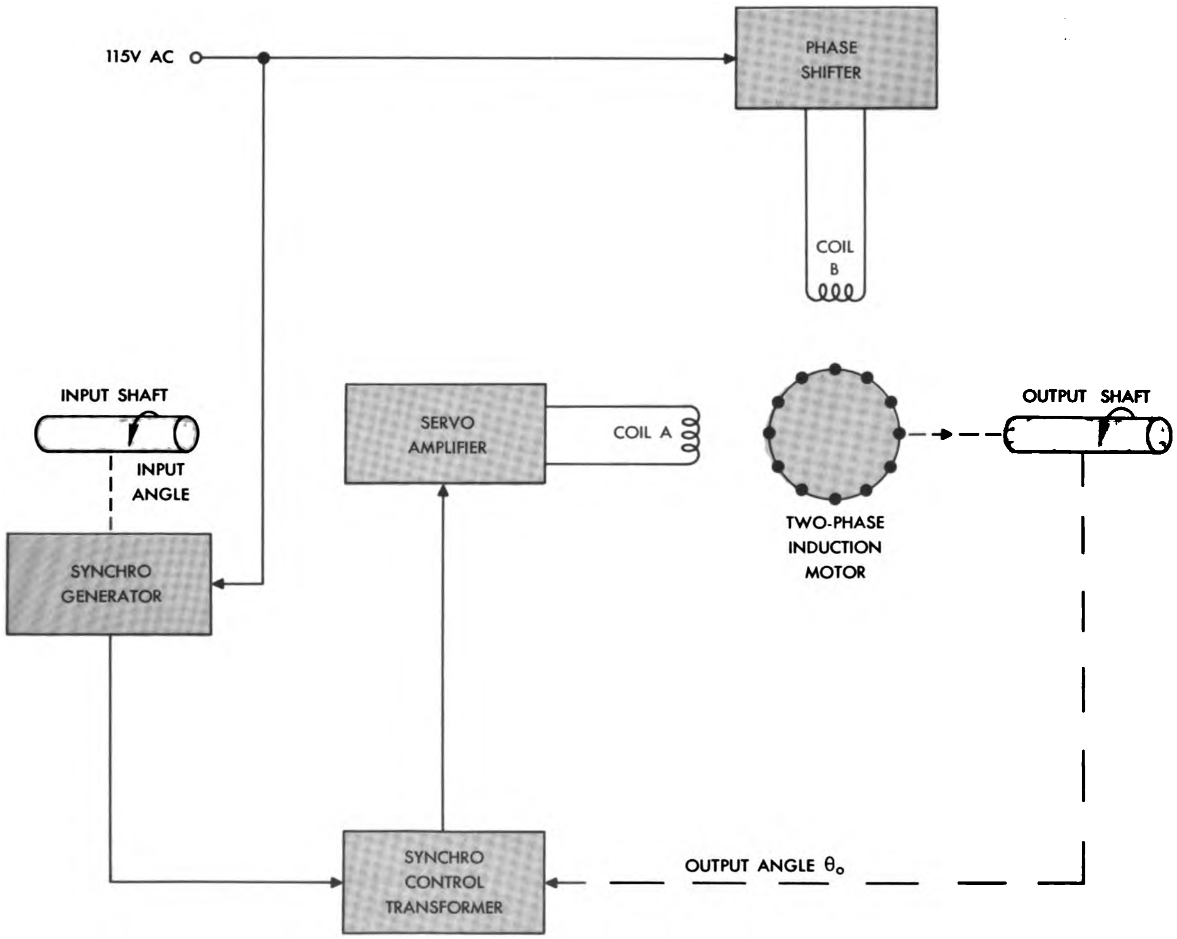
In the illustration shown above is a servomechanism which consists of a synchro generator, a control transformer, an amplifying device, and a servo motor. Mechanical connections are indicated by dashed lines.

As the rotor of the synchro generator is turned by the control shaft to which it is connected, an error voltage is developed across the rotor terminals of the control transformer. The phase and magnitude of this error voltage indicate the amount and direction of the turning of the control shaft. The error signal is amplified and applied to the servo motor. The motor turns in a direction and at a speed which is determined by the amplified error signal. As the motor turns, both the load and the rotor of the control transformer, which are mechanically linked to it, turn also. When the load is turned to the position which corresponds with that of the control shaft, the rotor of the control transformer is in the position where it nulls the error signal. With no error signal, the servo motor stops turning and the positions of the control and load shaft correspond.

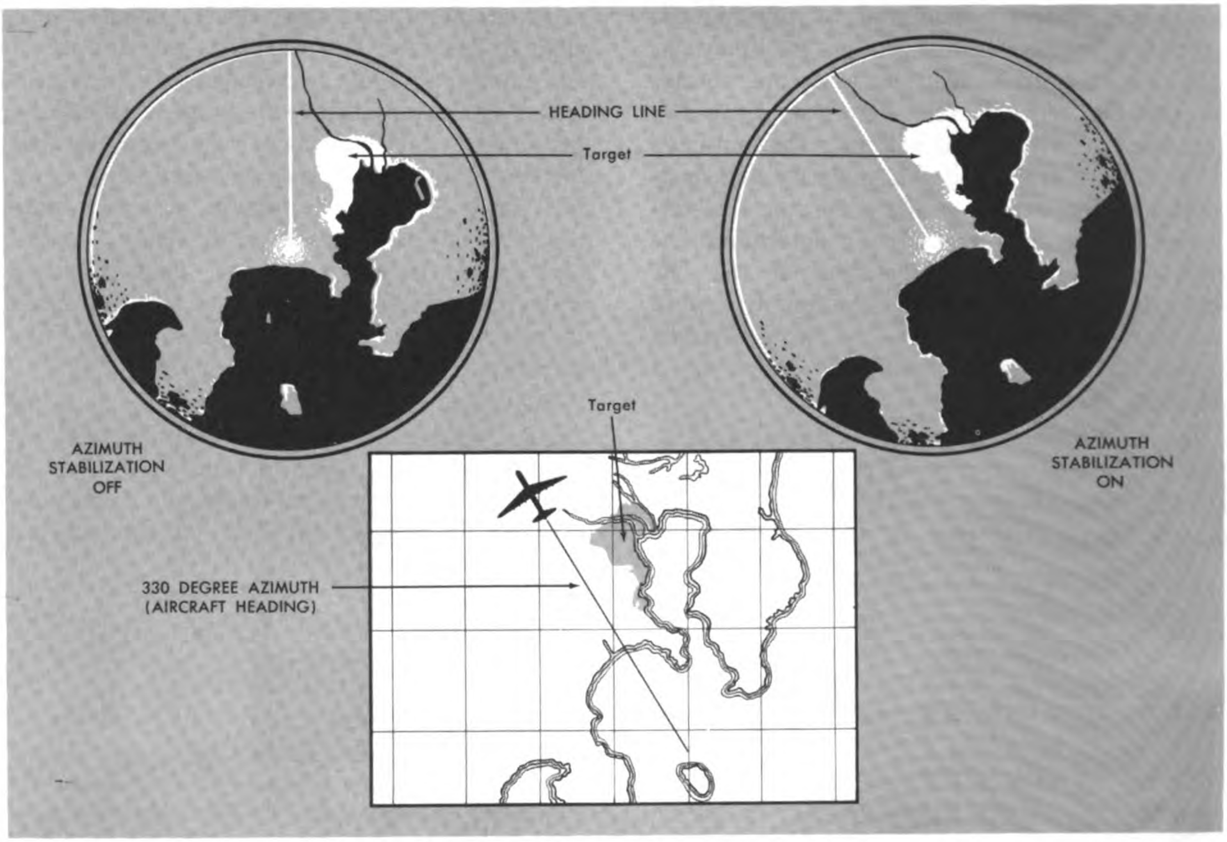
THE SERVO MOTOR. The servo motor of a servomechanism must be a motor whose direction and speed of rotation varies as the phase and magnitude of the error voltage changes. A DC motor or a two-phase AC induction motor may be used as a servo motor. If a DC motor is used, the control is achieved by varying the armature voltage. In a servo system known as a Ward Leonard Drive, the armature voltage for the DC servo motor is supplied by a DC generator to whose field windings the amplified error voltage is applied. The polarity and magnitude of the output of the generator is determined by the error voltage and determines the direction and speed of rotation of the DC servo motor. Such a system is commonly used in industry to control the speed of two generators coupled to the same shaft and driven by a constant speed motor.

In some instances, an amplidyne is used to replace the DC generator of the Ward Leonard Drive. An amplidyne is a DC generator which has an additional pair of brushes that are short-circuited. This short circuit makes it possible to generate a tremendous power output. An amplidyne is often considered to be a generator and an amplifier in one stage, thus making possible the development of very strong torques for controlling the motion of heavy equipment. DC servo motors are used in systems where much torque is required as in Ward Leonard and amplidyne servo systems. They are used when it is necessary to turn very large antennas of ground radar equipment or to control the direction of the firing of large guns.

Most low power servomechanisms used for smaller equipment employ two-phase AC induction motors as servo motors. In the system shown in the illustration below, the same voltage that is applied to R_1-R_2 of the synchro



Servomechanism with AC Servo Motor


HEADING LINE
Target
Target
AZIMUTH STABILIZATION OFF
AZIMUTH STABILIZATION ON
330 DEGREE AZIMUTH (AIRCRAFT HEADING)

Azimuth Stabilization

generator is shifted 90° in phase and applied to coil B of the servo motor. The amplified error voltage is applied to coil A of the servo motor. If there is no error voltage, the motor remains stationary since only one phase of voltage is being applied to the motor and no rotating magnetic field exists. If there is an error voltage, it is an AC voltage which is either in phase, or 180° out of phase with the voltage applied to R_1-R_2 of the synchro generator. Thus the voltage applied to coil A either leads or lags the voltage applied to coil B by 90°. A rotating magnetic field is created and the motor turns in the proper direction to null the error voltage. When the null position is reached, the motor stops. The speed with which the motor rotates at any instant depends upon the magnitude of the error voltage at that instant. Note that the motor speed is continuously decreasing until it becomes zero when the null position is reached.

A REVIEW OF IMPORTANT TERMS

A synchro is an AC machine which is used for the transmission of angular position information. There are three main types of synchros: ordinary synchros, synchro differentials, and control transformers.

Ordinary synchros may be either synchro generators or synchro motors. A synchro generator generates and transmits a torque. A synchro motor receives the torque transmitted by the generator and turns a light load in synchronization. An ordinary synchro system which consists of only a synchro generator and a synchro motor may be a complete system or merely form a basic component of a more intricate system.

A synchro differential may be either a generator or a motor althrough the generator is the more common. A differential generator combines an electrical signal from an ordinary synchro generator and a mechanical signal from a shaft and transmits the difference between the two positions to an ordinary synchro motor. A differential motor indicates the difference between signals from two ordinary synchro generators. The action of synchro differentials may be reversed so that a generator adds two signals and a motor indicates the sum of two signals.

A control transformer receives a signal from either an ordinary synchro generator or a differential generator and transmits a voltage signal called an error signal. The magnitude and phase of the error signal indicate how much and in which direction the rotor of the control transformer and that of the synchro generator are out of correspondence. Its output is amplified and then used to control how fast and in which direction the servo motor of a servomechanism will turn.

A servomechanism is a system which is used to synchronize the rotation of two shafts. However, a torque-amplifying device is added so that a small torque may be used to turn a heavy load from a remote position. A simple servomechanism consists of a synchro generator, a control transformer, an amplifying device, and a servo motor.

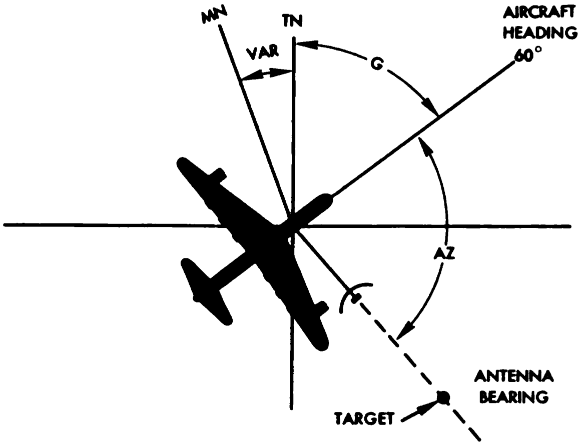
AN/APS-23 ANTENNA-INDICATOR SERVO SYSTEM

Up to now, most of the discussion of synchros and servomechanisms has been general. As had been pointed out previously, the radar antenna and the indicator sweep must be synchronized. When the basic synchro system is used with a radar, the scope presentation is always oriented with the aircraft, and the top of the indicator is the heading of the aircraft. The targets therefore change position on the scope as the aircraft changes heading. This results in a smearing of targets and also makes it necessary for the observer to reorient himself for each change of heading.

Features of Azimuth Stabilization

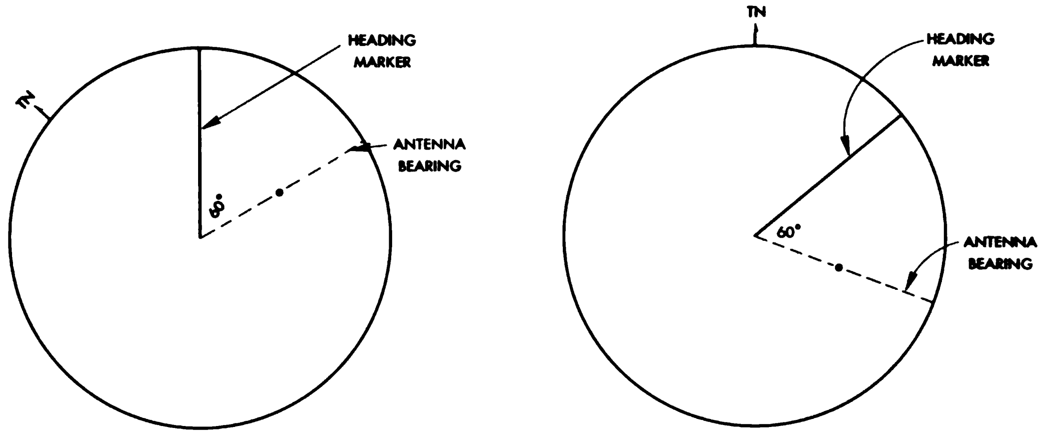
As you know, navigation charts are constructed with *true north* at the top. This is not only a custom, but is of practical value inasmuch as it allows anyone using the chart to orient himself quickly. A method of keeping the top of the scope always at true north so that it is unnecessary to twist and turn a chart to conform with the scope presentation has been devised. This system, known as *azimuth stabilization,* is incorporated in the AN/APS-23 to keep the top of the scope at true north, regardless of the heading of the aircraft. The effects of azimuth stabilization are shown in the illustration on the preceding page.

Another advantage of azimuth stabilization is that it enables you to read the true bearing of the target directly from the azimuth scale, and to plot the fix faster. Without azimuth stabilization, the bearing from the aircraft to the target, as read from the azimuth scale, is relative bearing and must be converted to true bearing before plotting by applying the formula, TB = TH + RB. You can see what factors actually affect azimuth stabilization in the illustration below. Notice that there are four factors involved — true north, magnetic north, aircraft heading, and antenna bearing.



Factors Affecting Azimuth Stabilization

If you assume that the aircraft and the antenna are stationary, and if you can cause the indicator selsyn to lead the antenna selsyn by the angle equal to true heading, then the top of the scope will be true north and the heading marker on the scope will indicate true heading. You can see this effect in the illustration at the top of page 494. Differential selsyns, which add or subtract the angles involved algebraically, are often used to accomplish azimuth stabilization. One input to the differential selsyn is the antenna bearing, and the other is TH supplied by the compass system.



WITHOUT AZIMUTH STABILIZATION

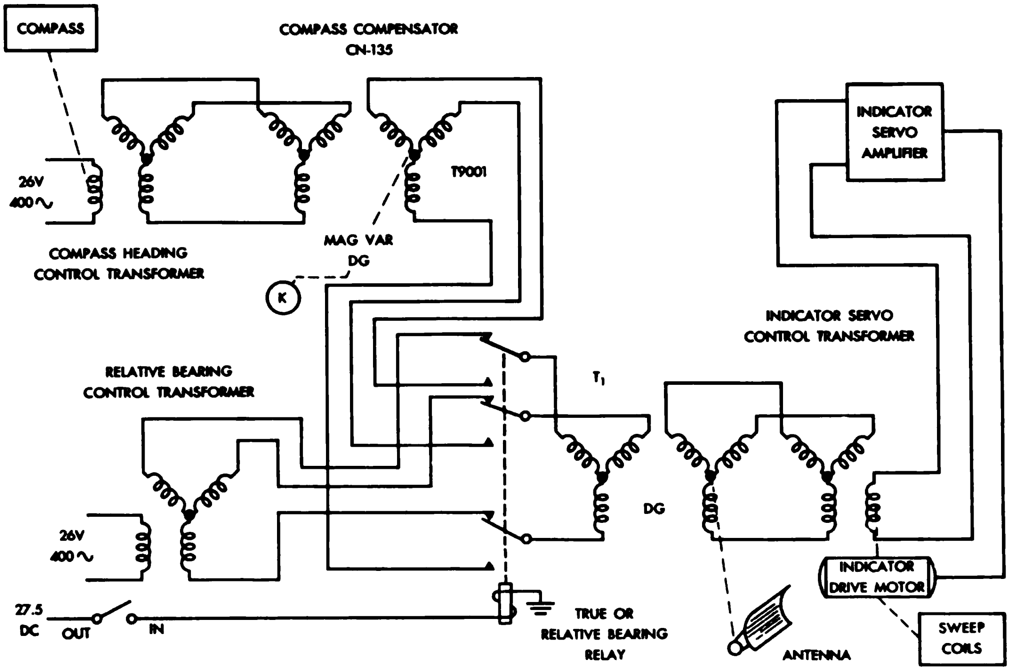
WITH AZIMUTH STABILIZATION

Orienting the Top of the Scope to True North

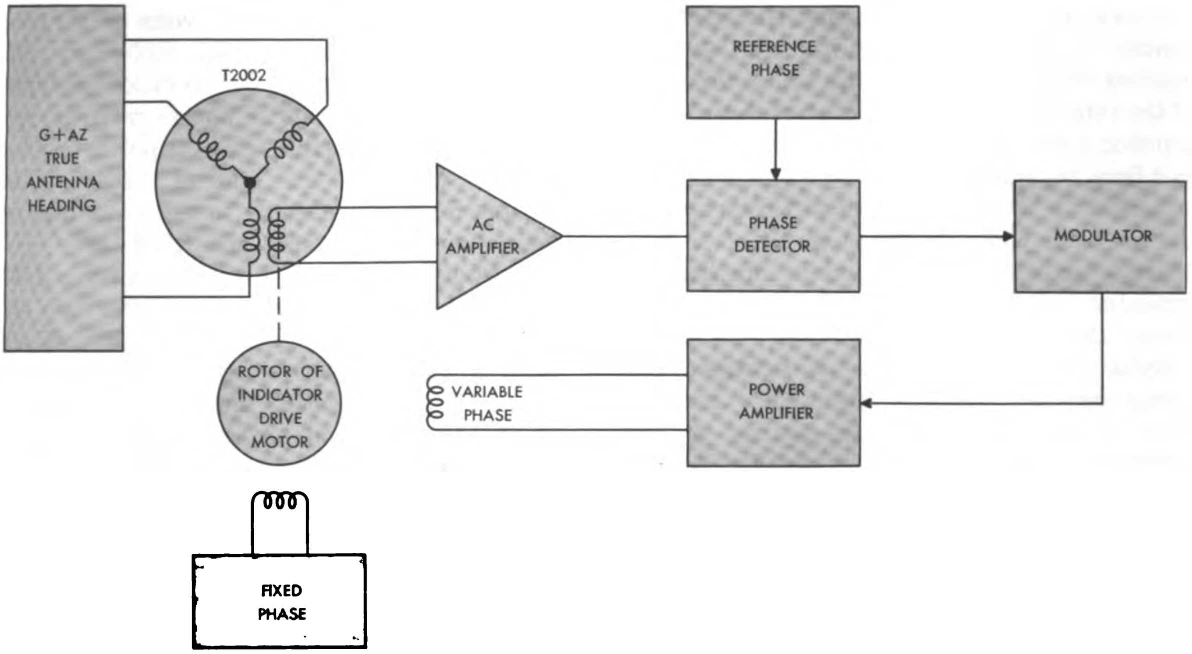
Block Diagram of AN/APS-23 Synchro System

In the illustration on page 472, you saw that a simple synchro system consisted of a synchro generator and a motor. This system is not particularly desirable for aircraft installations, because the low torque near the correct position causes inaccuracies and because the system is inflexible. The AN/APS-23 uses the synchro-servomechanism method to synchronize antenna and indicator and to provide azimuth stabilization. A general block diagram of this system is shown in the illustration below.

Notice that the information as to the heading of the aircraft is fed into the azimuth channel by the Gyrosyn compass system. The compass rotates the rotor of a control transformer, the output of which is compass heading. The com-



Block Diagram of AN/APS-23 Synchro System



Block Diagram of the Indicator AC Servo Loop

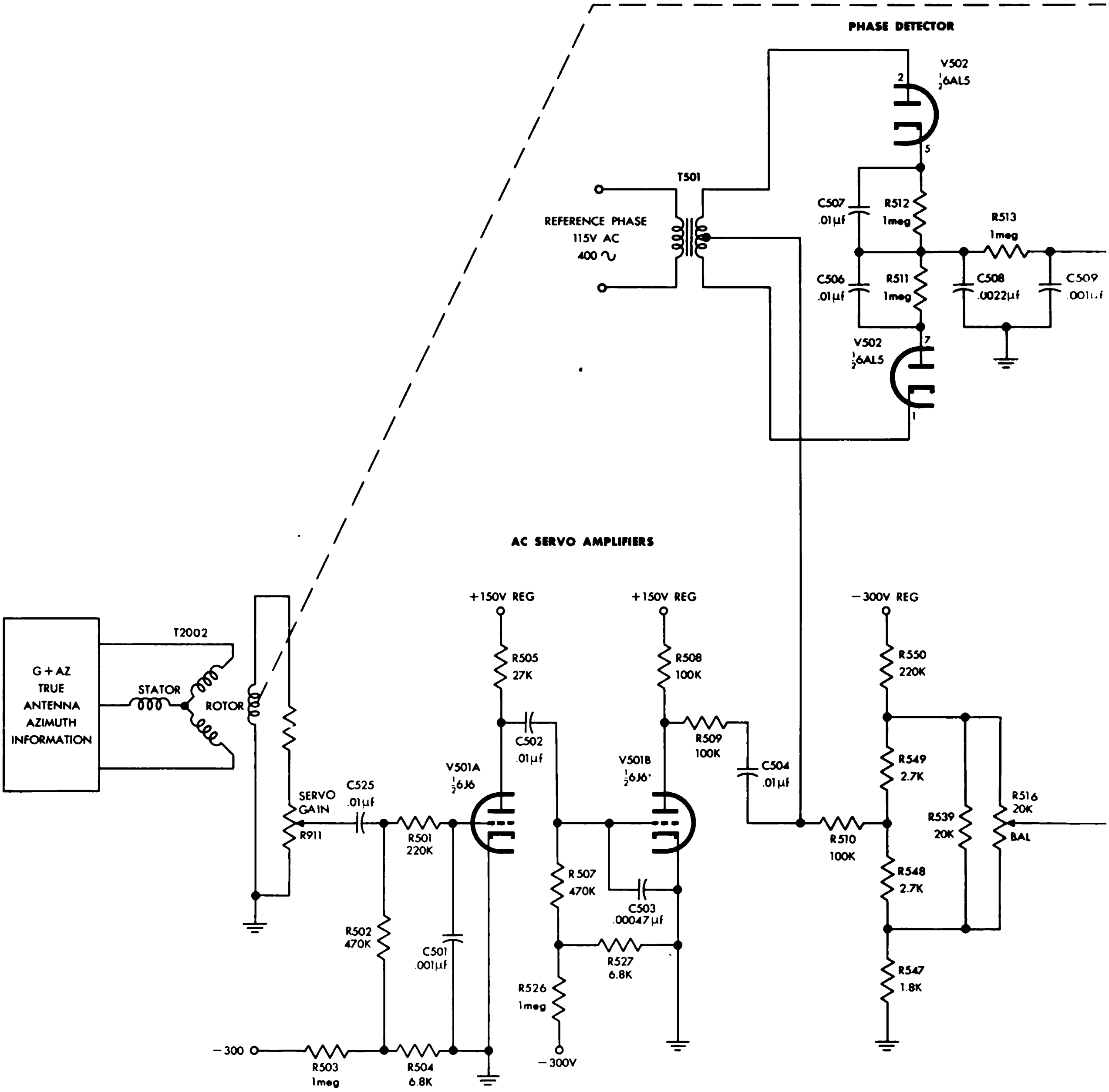
pass heading must be modified by a variation in order to obtain the true heading of the aircraft. This is accomplished by the Mag Variation Differential Generator, which enables the operator to add or subtract the number of degrees of variation from the compass heading. The output of the Mag Var DG is TH which is applied to the differential generator mounted on the antenna. The rotor of this differential generator is driven in synchronization with the antenna.

With the true bearing IN-OUT switch in the IN position, true heading information is applied to the stator of the differential generator T_1. The rotor of T_1 is driven by the antenna and supplies continuously changing information as to the azimuth of the antenna. This information is algebraically added to the TH and the result is applied to the stator of the Ind Servo CT as azimuth information. The voltage induced in the rotor of the Ind Servo CT is amplified by Ind Servo Amp and causes the Indicator Drive motor to rotate and drive the rotor to the null position. As the antenna continues to rotate, the Indicator Drive motor always has an input and continues to drive the rotor to a null. Since the sweep coils are connected to the shaft of the motor, they also turn and remain in synchronization with the antenna.

When the true bearing IN-OUT switch is in the OUT position, compass information is no longer fed into the azimuth chain. The relative bearing CT is used to give the longitudinal axis of the aircraft as a reference point to the azimuth chain. Now the top of the indicator becomes the heading of the aircraft instead of true north, and relative bearings are read from the azimuth ring.

Indicator AC Servo Loop

Generally, there isn't anything new in the azimuth chain other than the Indicator AC Servo system. Up to the point of actually driving the sweep coil, the synchro system has not yet had to accomplish any work. It has only had to provide azimuth information. Since the synchros used are low torque devices, they cannot drive the sweep coil directly without introducing a considerable error which would cause the sweep to lag the actual antenna position by many degrees. To prevent this lag, a method was devised which uses the azimuth information to control another device which, in turn, drives the sweep coil. The method employed, is the AC Servo Loop which is shown in block diagram in the illustration above.

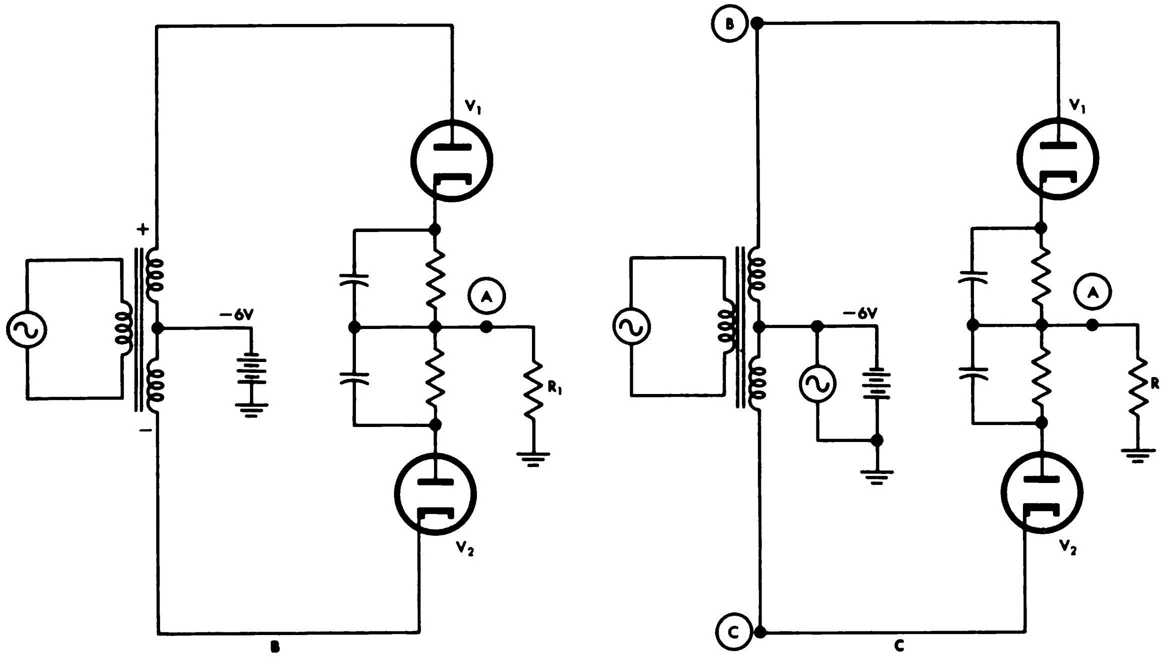

PHASE DETECTOR
V502
½6AL5
2
5
T501
REFERENCE PHASE
115V AC
400 ∿
C507
.01μf
R512
1meg
R513
1meg
C506
.01μf
R511
1meg
C508
.0022μf
C509
.001μf
V502
½6AL5
7
1
AC SERVO AMPLIFIERS
+150V REG
+150V REG
−300V REG
T2002
G + AZ
TRUE
ANTENNA
AZIMUTH
INFORMATION
STATOR
ROTOR
R505
27K
R508
100K
R550
220K
C502
.01μf
R509
100K
V501A
½6J6
V501B
½6J6
C504
.01μf
R549
2.7K
R516
20K
SERVO
GAIN
R911
C525
.01μf
R501
220K
R510
100K
R539
20K
BAL
R507
470K
C503
.00047μf
R548
2.7K
R502
470K
C501
.001μf
R527
6.8K
R526
1meg
R547
1.8K
−300
R503
1meg
R504
6.8K
−300V

If a bias voltage is applied to the center tap of the transformer, as shown in diagram B below, the currents in the common resistor R_1 will not cancel and point A will be at a potential equal to the bias applied to the center tap.

Consider the AC signal applied to the primary of the transformer the reference phase and apply an AC signal of unknown phase to the center tap of the transformer *in addition* to the -6 volts bias. You can tell whether the unknown AC signal is in phase or out of phase with the reference voltage by determining whether the voltage across R_1 is above or below the -6 volts level. Why this is so can be seen from diagram C. If the unknown AC signal is in phase with the reference voltage and is in phase with point B, (going positive as B goes positive) then V_1 will cause more current to flow through R_1 than V_2. This causes the potential at point A to decrease in a negative direction. If the unknown AC signal is in phase with point C, V_2 conducts more and the voltage across R_1 increases in a negative direction. You can therefore determine whether the applied AC signal is in or out of phase with the reference voltage by determining whether point A is above or below the bias level.

The phase detector in the Indicator Servo Loop illustration on page 499 works precisely in the manner just described. The voltage applied to the primary of transformer T501 is 115V, 400 ~ AC and is called the reference voltage. It applies voltages of equal amplitude but of opposite polarity to the plate 2 of one half and cathode 1 of the other half of V502. The tubes conduct equally, within the limits of manufacturing tolerances, and the current cancels out in the RC filter composed of R513, C508, and C509. The -6 volts bias applied to the center tap of T501, however, causes the filter to charge to a potential of -6 volts. The amplified signal from the rotor of T2002 is also applied to the center tap of T501 through the coupling condenser C504. The result is that the DC voltage across the filter is either more or less negative than the negative bias introduced at the center tap of T501, depending upon whether the voltage from the rotor is in phase or out of phase with the reference voltage. This output voltage is then fed to the modulator whose function it is to produce a 400 cycle output, the phase of which is determined by the phase detector.

Modulator. Let us now consider the basic modulator circuit shown on page 503.



B

C

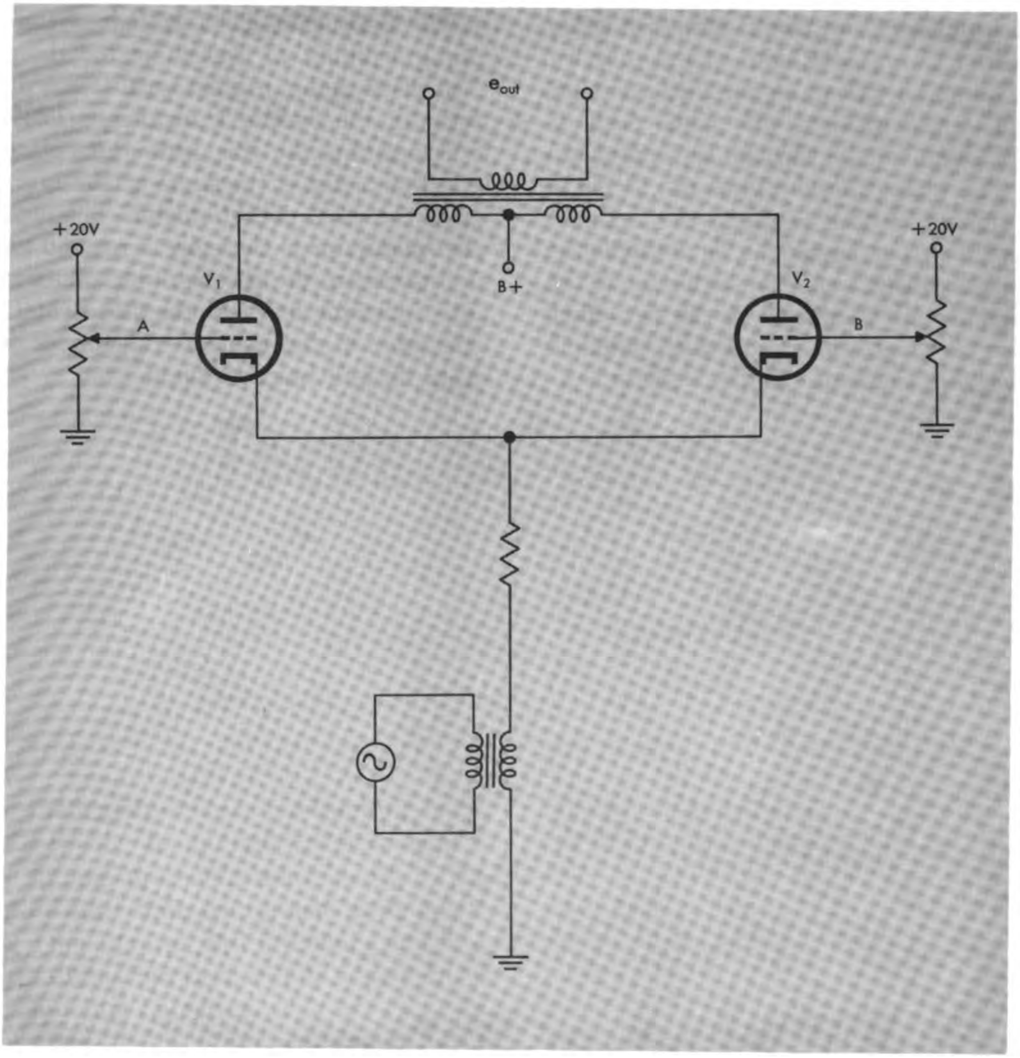
Phase Detector — Unequal Inputs

Assume that grid B is at ground potential and grid A is at plus 20 volts. Under this condition if a negative signal is applied to the common cathode, V_1 will conduct and V_2 will cut off. At the same time, the plate of V_1 will go negative and will be in phase with the applied signal. Similarly, if grid A is grounded and grid B is at plus 20 volts, a negative signal applied to the common cathode will cause V_2 to conduct and V_1 to cut off. The important thing to observe is that the plates are connected by a push pull output transformer. If a current flows from the plate of V_1 to the center tap and later from V_2 to the center tap, the output phase will be reversed.

The modulator in the schematic of the Indicator Servo Loop on page 499 works in much the same way, since its output phase is determined by which of the plates is conducting more. The output of the phase detector, which is above or below its −6 volt reference, determines which plate conducts more. Transformer T507 steps down the 115 V 400 ∼ input voltage and applies this voltage in push-pull to the cathodes of V504 and V505. In addition to this, a bias voltage is applied to the cathode through R518 and R519. Grid 6 of V504 and grid 5 of V505 are connected to the output of the phase detector while the other grids are connected to a bias potentiometer R516, Servo Bal. The Servo Bal makes it possible to adjust this bias above or below −6 volts to compare to the reference level bias of the phase detector. Notice also that one side of the push-pull output transformer T502 is connected to plate 1 of V504 and V505, while the other side is connected to plate 2 of V504 and V505.

ROTOR AT NULL POSITION. Assume first that the rotor of T2002 in the Indicator Servo Loop schematic is at a null position and the output of the phase detector is −6 volts. Resistor R516 is adjusted until the DC voltage on grid 5 of V504 and grid 6 of V505 is the same as the output of the phase detector (−6 volts) which is fed to the other pair of grids. For any given half cycle of voltage impressed on the cathodes by T507, the cathodes are 180° out of phase with each other. Since grids 5 and 6 of V504 are at the same potential, the plate currents will increase and decrease in phase; however, they will be 180° out of phase with the plate currents of V505 because the voltages on the cathodes are 180° out of phase. The two plate 1 currents flow through the same half of transformer T502. Because they are 180° out of phase, the cathode voltage variations cancel out, causing a steady DC flow. Likewise, the two plate currents flow through the other half of the transformer. Since these currents, too, are 180° out of phase, the cathode variations also cancel out, leaving a DC flow. These direct currents flow in opposing directions to the center tap of T502 with the result that there is zero net flux and no output from the transformer.

ROTOR WITH PHASE DETECTOR MORE NEGATIVE. Assume that the rotor is not at a null position, and the phase is such that the output of the phase detector is more negative. Grid 6 of V504 and grid 5 of V505 become more negative, while grid 5 of V504 and grid 6 of V505 remain at −6 volts as before. Consider the first half cycle of the AC signal applied to the cathodes, and assume that the cathode of V505 is going positive while the cathode of V504 is going negative. During this time, the current through V504 is increasing because the signal on the cathode is decreasing. Plate 2 conducts as it did during the balanced condition while the current through plate 1 is reduced from the balanced condition because of the increase in bias felt on the grid. At the same time, the current through V505 is decreasing because of the increasing signal applied to its cathode. Plate 1 conducts as it did during the balanced condition, while the conduction through plate 2 is reduced because of the increased bias. The net result is that there is reduced current flow through section 1-2 of the transformer and increased current flow through section 2-3. The magnetic fluxes created do not cancel out and a voltage is induced in the secondary of T502. Note that during the second half cycle the flux is in opposite direction through transformer T502 and that the polarity of the output is reversed. However, this reversal occurs in phase with the reference voltage and the signal injection voltage. When the output of T502 is amplified and applied to the indicator drive motor, the direction of rotation remains the same, since the reference voltage connected to the fixed phase winding changes at the same



Basic Modulator Circuit

time. The reference voltage applied to the fixed phase winding has been shifted 90° by a phase shifting network. Otherwise, the voltages on the two windings would either be in or out of phase, and the motor would not rotate.

ROTOR WITH PHASE DETECTOR LESS NEGATIVE. If the phase of the rotor voltage causes the output of the phase detector to go less negative, grid 6 of V504 and grid 5 of V505 would feel a decrease in bias. Consider the first half cycle of operation again with the cathode of V504 going negative and the cathode of V505 going positive. The current through V504 is increasing due to the decreasing signal on the cathode. Plate 2 is conducting as it did during the balanced condition while the current through plate 1 is increasing, due to the decreased bias on grid 6. Simultaneously the current through V505 is decreasing due to the increasing signal on its cathode. Plate 1 conducts as it did in the balanced condition while plate 2 conducts more heavily due to its decreased bias. The overall effect is that the current flow through section 1-2 of T502 is more than through section 2-3. The magnetic fluxes do not cancel and a voltage is induced in the secondary of T502. The output of T502 is reversed in polarity to the output obtained under the previous conditions when the phase detector output went more negative.

ROTATING THE INDICATOR DRIVE MOTOR. The output of T502 is applied in push-pull to the grids of the power amplifier tubes V506 and

V507. The plates of the power amplifier tubes are connected to opposite sides of the primary of T503. Unfiltered, full-wave, rectified reference voltage of approximately 440 volts peak is supplied as plate voltage through the center tap of T503. The plate voltage pulses are in phase with the signals on both grids. In this manner, greater useful output power can be obtained for a given plate dissipation. The cathodes of the power amplifier are connected to the secondary of T503 so as to produce a negative feedback across R523, the common cathode resistor.

The output of T503 is fed to the variable phase winding of the indicator drive motor. The drive motor rotates until it has driven the rotor of the control transformer T2002 to the null position. At this time the output of the phase detector is the same as the bias supplied from the BAL potentiometer. The output from the modulator is therefore zero, and no voltage is impressed on the variable phase winding. The indicator drive motor stops rotating until the rotor of the control transformer is shifted from its null position again.

Chapter 40

Technical Orders

One of the main purposes for teaching basic electronic theory to navigators is to enable them to become proficient more easily in the use of new electronic equipment as it is developed. This chapter will introduce you to a technical publication that will be very helpful in achieving this objective.

Every piece of Air Force equipment has a publication that describes its operation, maintenance, service instructions, etc. These publications are Technical Orders (or simply TO's); a complete set of them will be located at each Air Force Base for the equipment at that base. Each book may be thought of as the "Bible" for a given piece of equipment.

Technical Order Index

The TO library is indexed in a manner similar to a public library. In order to find the TO for a particular piece of equipment, you should know how to use the indexing system.

There are TO indexes for each category of Air Force equipment. As a navigator, you will be primarily interested in the Numerical Index of "Airborne Electronic Equipment Publications" (TO 0–1–12). A detailed breakdown of a typical TO number will best explain the indexing system used in TO 0–1–12. We will use the number 12P6–2APS42–21 and explain each portion in turn. This number consists of seven parts:

12 P 6 –2 APS 42 –21

First, the *12* stands for the category of the equipment, in this case, "Airborne Electronic Equipment." If the equipment were "Propellers," the number would be *3*; "Fuel and Oils" would be *37*, and "Missiles" would be *21*.

The next part of the number is alphabetical and designates the major group to which the equipment belongs, or what type of equipment it is. *P* stands for radar systems. Similarly, *M* is for meteorological, *W* is for weapons, *F* is for fire control systems, and *R* is for radio systems.

The *6* indicates that this is primarily a search and height-finding radar. A *5* would indicate navigational equipment; a *4* would indicate IFF equipment, etc.

The next number indicates the type of nomenclature used. The *–2* means that "AN" nomenclature is used; a *–3* indicates Signal Corps nomenclature, and a *–5* indicates Air Force nomenclature.

The next part of the number, in this case *42*, identifies the specific Air Force radar search set.

The last part of the number identifies the individual TO publication. Numbers ending in *-1*, such as 1, 11, 21, 31, up to 491, and then 1001, 1011, etc., are *Operating Instructions*. The *-2's* (2, 12, 22, 32, etc.) are *Service* or *Maintenance Instructions;* the *-4's* are *Parts Catalogs*, and the *-6's* are *Preventive Maintenance Instructions*. The specific TO we are dealing with, *-21*, contains operating instructions for the AN/APS-42.

The Technical Order indexing system just described is relatively new. For some TO's, the old system may still be in effect. Where a TO carries the old designation, its new number may be found in TO 0-4-1, which is a cross-reference table of old to new numbers.

Technical Order Contents

Usually there are eight sections in a TO. Section I contains general information about the equipment. Section II contains installation and adjustment instructions. Section III contains operating instructions. Section IV contains material on the theory of operation. Section V contains maintenance instructions. Section VI contains supplementary data. Section VII contains a table of replaceable parts. Section VIII contains drawings, wiring diagrams, and schematics. (This layout is general and many TO's may differ from it.)

Eventually, basic circuit theory will be written in a TO such as 31-1-141. Then TO's on specific equipment will not need to cover fundamental circuits and there need not be so many volumes, in the *-2's* especially.

Form 781 Write-Ups

The amount and complexity of airborne electronic equipment being what it is, it would be difficult indeed to estimate the number of manhours spent in repairing or maintaining this equipment. However, aircraft crew members themselves can do much to keep maintenance time at an absolute minimum simply by making their malfunction write-ups in the aircraft Flight Report and Maintenance Record, Form 781, as complete and accurate as possible.

Location of the specific source of trouble is probably the most time-consuming phase of repair. It is in this area that the write-up made by the crew member can be of greatest value. A complete write-up should contain an exact description of the indication of trouble ("PPI blank," "no MOD HV," etc.) and a record of "in-flight" corrections attempted.

The importance of making good write-ups cannot be overemphasized. The amount of ground-crew troubleshooting thus eliminated would result in a saving of millions of dollars to the government.

Chapter 41

Mathematics for Navigators

As a navigator, you will find a knowledge of mathematics of great value to you in the understanding of electronic and radar circuits. This chapter presents in compact form the mathematics that applies especially to the field of electronic fundamentals and radar. In it, you will find a review of algebra — the fundamental processes, the solution of equations, special products and factoring, exponents, radicals, and complex numbers. As an aid in the solution of alternating-current problems, trigonometric functions, logarithms, and vectors are presented. You will find also a brief discussion of how to use a slide rule to perform computations. The use of a DR computer, both as a slide rule and as an aid in solving AC problems, is presented with illustrative examples. Although the discussion in this chapter is necessarily brief, you will find that it covers the mathematics you will need in electronics and radar. If you require a more complete explanation of any of the mathematics covered, refer to a standard text on that subject.

TERMINOLOGY OF ALGEBRA

An *algebraic expression* is any combination of signs, numerals, and letters used for numbers, written according to the rules of algebra. For example, $a + b$, l, lw, and $-\frac{x}{2y}$ are algebraic expressions. Numbers represented by letters are called *literal* numbers. Thus a, b, l, w, x, and y are literal numbers. Signs may indicate whether the numbers are positive or negative, or they may indicate operations, for example, addition, subtraction, multiplication, division, a root, or a power.

Multiplication of two algebraic numbers need not be indicated by the symbols for multiplication — the times sign (×) or the dot (·). Just putting them together indicates multiplication. Thus, the product of a and b can be written ab. In the expression ab, a and b are called *factors* of the product. Likewise 6, a, and b are factors of $6ab$.

Each factor of a product is the *coefficient* of the other factors. In the algebraic expression $6ab$, 6 is the coefficient of ab, $6a$ of b, and $6b$ of a. In this expression, 6 is called the numerical coefficient and b the literal coefficient of a. The word coefficient is usually restricted to mean only the numerical or arithmetic coefficient.

The size or magnitude of a number is called its *absolute value*. Absolute value refers only to the magnitude with no consideration to the sign preceding the number. For example, the absolute value of $+7$ is 7, and the absolute value of -9 is 9. Both $+3$ and -3 have the same absolute value, 3.

An *exponent* is a number or letter which indicates the power to which a quantity is to be raised, or the number of times the quantity (called the base) is multiplied by itself. For example, 4^3 is read *4 to the third power*, and means $4 \times 4 \times 4$. The expression e^x is read *e to the x power*, and means that e is multiplied by itself x times.

Any arithmetic or literal number, or the product or the quotient of these numbers, is called a *term*. For example $4, x, y, 25b, ab, \frac{x}{y}, \frac{3a^2b}{c}$ are terms.

Terms that have the same literal parts are called *similar* or *like* terms. Thus a, $7a$, and $12a$ are like terms. Terms, such as x^2, y, lh, which do not have the same literal parts, are called *unlike* terms.

You can add or subtract two or more numbers if they are like terms. For example, you can add $4x$ to $5x$ to produce the sum $9x$. Likewise, you can add $5x^2$ to $10x^2$, since the literal parts x^2 are alike. (Literal numbers raised to the same power are like terms.) The sum of numbers such as $4x$ and $5m^2$ must be indicated as $4x + 5m^2$ since x and m^2 are not like terms.

An algebraic expression containing two or more terms joined together by a plus or minus sign is called a *polynomial*. Thus, the expressions $a + b - c$ and $ab - ac + \frac{a}{c}$ are polynomials. A polynomial with just two terms such as $5x + y$ is a *binomial*. A polynomial with only three terms is a *trinomial*. The expression $5x^2 + ab + c$ is a trinomial. A single term, such as a, x, $10x^2y$, is a monomial.

SOME RULES OF ALGEBRA

Addition

To add two numbers with the same sign, add their absolute values and prefix the common sign.

Thus, $+6$ *added to* $+3$ *equals* $+9$
-6 *added to* -3 *equals* -9
$+6a^2$ *added to* $+3a^2$ *equals* $+9a^2$
$-6m^2n$ *added to* $-3m^2n$ *equals* $-9m^2n$

To add two numbers with opposite signs, take the difference between their absolute values and prefix the sign of the number with the larger absolute value.

Thus, $+6$ *added to* -3 *equals* $+3$
-6 *added to* $+3$ *equals* -3
$+6a^2$ *added to* $-3a^2$ *equals* $+3a^2$
$-3m^2n$ *added to* $+6m^2n$ *equals* $+3m^2n$

Subtraction

The process of subtraction is opposite to the process of addition. To subtract a quantity from another, change the sign of the quantity to be subtracted, then add the quantities, following the rules of addition.

For example

$$(-6)-(-3)=(-6)+(+3)=-3$$
$$(-3)-(-6)=(-3)+(+6)=+3$$
$$(+6)-(-3)=(+6)+(+3)=+9$$
$$(+6)-(+3)=(+6)+(-3)=+3$$

Signs of Grouping

Certain symbols, called *signs of grouping*, tie or group together several quantities affected by the same operation. The most commonly used symbols of grouping are parentheses ( ), brackets [ ], braces { }, and the vinculum $\overline{\qquad}$. All these symbols indicate that the quantities affected by them are to be treated as single quantities. For example, the expression $(4a^2 - 3ab) - (3a^2 + 2ab)$ means that the quantity $3a^2 + 2ab$ is to be subtracted from $4a^2 - 3ab$. To perform this operation, change the signs of the expression $3a^2 + 2ab$, and add algebraically,

$$\begin{array}{r} 4a^2-3ab \\ \underline{-3a^2-2ab} \\ a^2-5ab \end{array}$$

In removing a symbol of grouping preceded by a minus sign, change the signs of all terms included by the symbol. When one symbol of grouping occurs within another, remove the innermost symbol first. You may remove both symbols at the same time if you are quite careful, but this practice can easily lead to errors of sign. Therefore, it is best to remove only one symbol at a time. The following example illustrates the procedure for removing a symbol of grouping occurring within another:

$$\begin{aligned}
&3a^2-[2am-(2a^2+5am)+a^2]\\
&3a^2-[2am-2a^2-5am+a^2]\\
&3a^2-2am+2a^2+5am-a^2\\
&4a^2+3am
\end{aligned}$$

Multiplication

In the multiplication of algebraic terms, there are three things to consider — the *sign*, the *exponent*, and the *coefficient*.

THE SIGN. The product of two terms with like signs is positive. The product of two terms with unlike signs is negative.

THE EXPONENT. The exponent of any letter in the product is the sum of the exponents of the factors with the same base.

Thus, $x^2 \times x^3 = x^{(2+3)}$, *or* x^5

Likewise, $y^3 \times y^7 = y^{10}$

THE COEFFICIENT. The arithmetic coefficient of the product is the product of the absolute values of the coefficients of the terms being multiplied.

Thus, $6x^2 \times 2x^3 = 12x^5$

Likewise, $3y^3 \times 3y^7 = 9y^{10}$

Examples

$$\begin{aligned}
&(3a^2)\times(5a^3)=15a^5\\
&(2ax)\times(-3a^2xy)=-6a^3x^2y\\
&(m^2n)\times(-7mn^2)=-7m^3n^3\\
&(4xy)\times(-xy^3)=-4x^2y^4\\
&(-a)\times(-5a^3)=5a^4
\end{aligned}$$

A literal quantity with no written exponents means the quantity is to the first power — that is, it has an exponent of one.

Multiplication of polynomials is similar to the multiplication of numbers consisting of several digits in arithmetic. Multiply each term of the multiplier by every term of the multiplicand and add the partial products. Thus, in multiplying $2e^2 -3e -5$ by $-4e +7$, proceed as follows:

$$\begin{array}{rrrr}
 & 2e^2 & -3e & -5\\
 & & -4e & +7\\
\hline
-8e^3 & +12e^2 & +20e & \\
 & +14e^2 & -21e & -35\\
\hline
-8e^3 & +26e^2 & -e & -35
\end{array}$$

Division

In the division of one term by another, the *sign*, the *coefficient*, and the *exponent* must be considered in obtaining the quotient.

THE SIGN. The quotient of two positive or two negative quantities is positive; the quotient of a positive and a negative quantity is negative.

THE EXPONENT. The quotient of powers having like bases has the same base as the given powers, and an exponent which is the difference of the exponents of the given powers.

Thus, $\frac{x^4}{x^2}=x^2$

Likewise, $\frac{x^{10}}{x^3}=x^7$

THE COEFFICIENT. To obtain the coefficient of the quotient, divide the absolute value of the coefficient of the dividend by the absolute value of the coefficient of the divisor.

Thus, $\frac{6x^4}{3x^2}=2x^2$

Likewise, $\frac{8x^{10}}{2x^3}=4x^7$

ZERO AND NEGATIVE EXPONENTS. In division there is the possibility of zero and negative exponents, as the following examples indicate.

$$\frac{a^5}{a^5}=a^{(5-5)} \text{ or } a^0, \text{ or } \frac{a^4}{a^5}=a^{-1}, \frac{a^3}{a^5}=a^{-2}$$

Any quantity with a zero exponent is equal to *one.*

Thus, $\frac{a^5}{a^5}=a^0=1$ *or* $\frac{x^3}{x^3}=x^0=1$

Any quantity with a negative exponent is equal to the reciprocal of that quantity with the corresponding positive exponent.

Thus, $a^{-2}=\frac{1}{a^2}$ *or* $\frac{1}{x^{-3}}=x^3$

Any factor can be moved from the numerator to the denominator of a fraction or vice versa without changing the value of the fraction, if the sign of the exponent is changed.

Thus, $\frac{3x^2y}{a^3b}=\frac{3x^2ya^{-3}}{b}$

DIVISION OF POLYNOMIALS. Division of one polynomial by another is similar to long division in arithmetic. One difference is that the dividend, divisor, and the remainder (if there is one) must be arranged in order of ascending or descending powers of some letter.

Example

Divide $30c^4+3-82c^2-5c+11c^3$ by $3c^2-4+2c$

Solution

$$
\begin{array}{rl}
 & 10c^2-3c-12 \\
3c^2+2c-4 & \overline{)\,30c^4+11c^3-82c^2-5c+3} \\
 & \underline{30c^4+20c^3-40c^2} \\
 & -9c^3-42c^2-5c \\
 & \underline{-9c^3-6c^2+12c} \\
 & -36c^2-17c+3 \\
 & \underline{-36c^2-24c+48} \\
 & 7c-45 \text{ remainder}
\end{array}
$$

SOLUTION OF EQUATIONS

An equation is a statement that two quantities are equal. To solve an equation means to find the value or values of the literal numbers for which the equation holds true. In solving an equation, you will have to use *axioms*, statements accepted as true without proof. Here are a few of the more commonly used axioms:

1. If the same number is added to or subtracted from both sides of an equation, the result is still an equation.
2. If both sides of an equation are multiplied or divided by the same quantity (not zero), the result is still an equation.
3. If like roots or powers are taken of both sides of an equation, the result is still an equation.

Other processes which are based upon the axioms and are often applied in the solution of equations are *transposition, changing signs*, and *cancellation*. Transposition is the transferring of a term from one side of an equation to the other and changing its sign. This is merely another way of saying that the term was added to or subtracted from both sides (axiom 1). Changing signs of terms on both sides of an equation is merely another way of multiplying or dividing both sides by -1 (axiom 2). Cancellation is the means of collecting terms or of transposing and collecting terms.

The examples which follow illustrate the use of axioms in the solution of equations.

Examples

1. Given $x-5=3$, find x.

Solution

Add 5 to both sides of the equation (axiom 1),

$$x-5+5=3+5$$

Then collect terms.

$$x=8$$

2. Given $5x-4=21$, find x.

Solution

Add 4 to both sides of the equation (axiom 1),

$$5x-4+4=21+4$$

Then collect terms on both sides,

$$5x=25$$

Divide both sides by 5 (axiom 2),

$$x=5$$

3. Given $\frac{1}{3}x+5=8$, find x.

Solution

Subtract 5 from both sides of the equation (axiom 1),

$$\frac{1}{3}x+5-5=8-5$$

Collect terms,

$$\frac{1}{3}x=3$$

Then multiply both sides of the equation by 3 (axiom 2),

$$x=9$$

4. Given $\frac{4}{5}x+5=25-1\frac{1}{5}x$, find x.

Solution

Subtract 5 from both sides of the equation (axiom 1),

$$\frac{4}{5}x=20-1\frac{1}{5}x$$

Then add $1\frac{1}{5}x$ to both sides (axiom 1),

$$\frac{4}{5}x+1\frac{1}{5}x=20-1\frac{1}{5}x+1\frac{1}{5}x$$

Collect terms,

$$2x=20$$

Divide both sides by 2 (axiom 2),

$$x=10$$

5. Given $16-5(x+3)=4(2x+1)-9\frac{1}{2}$, find x.

Solution

Remove the parentheses,

$$16-5x-15=8x+4-9\frac{1}{2}$$

Collect terms,

$$1-5x=8x-5\frac{1}{2}$$

Subtract 1 from both sides (axiom 1),

$$-5x=8x-6\frac{1}{2}$$

Subtract 8x from both sides (axiom 1),

$$-13x = -6\frac{1}{2}$$

Then divide both sides by −13 (axiom 2),

$$x = \frac{1}{2}$$

6. Given $\sqrt{x} + 2 = 4$, *find x.*

Solution

Subtract 2 from both sides of the equation,

$$\sqrt{x} = 2$$

Then square both sides (axiom 3),

$$x = 2^2 \text{ or } 4$$

7. Given $z^2 = r^2 + x^2$, *find r.*

Solution

Subtract x^2 *from both sides of the equation,*

$$z^2 - x^2 = r^2$$

Take the square root of both sides of the equation (axiom 3),

$$\sqrt{z^2 - x^2} = r$$

MONOMIAL SQUARE AND SQUARE ROOTS

When a monomial is multiplied by itself, the product is the *square* of the coefficient multiplied by each literal quantity with its exponent doubled.

The square root of a monomial is the square root of the coefficient times each literal number with its exponent halved. There are two roots—a negative root and a positive root. This is true because the sign of the product of two negative or two positive quantities is positive.

Examples

$$(3a^2)^2 = 9a^4$$

$$(-5mn^3)^2 = 25m^2n^6$$

$\sqrt{64a^4} = \pm 8a^2$ *(The symbol ± means that it could be either* $+8a^2$ *or* $-8a^2$*)*

$$\sqrt{50a^3} = \pm 7.07a^{1.5}$$

SPECIAL PRODUCTS AND FACTORING

Factoring is the process of finding two or more quantities, each called a factor, whose product is equal to a given quantity.

Common Monomial Factor

The most simple form of polynomial factoring is finding the common monomial factor. The expression $5x + 5y$ is a polynomial in which 5 is a factor common to both monomials. Factor this expression by dividing $5x + 5y$ by 5. Write the result as $5(x + y)$.

Examples

$$ax + ab - az = a(x + b - z)$$

$$4mx - 8my + 2mz = 2m(2x - 4y + z)$$

$$5a^2b - 10a^2b^2 - 15ab^3 = 5ab(a^2 - 2ab - 3b^2)$$

Product of Two Binomials with Common Term

A frequently encountered product has the form $(x + a)\ (x + b)$, which, when multiplied, gives $x^2 + x\ (a + b) + ab$. This relationship can be stated as a rule: The product of two binomials having a common term is the square of the common term, plus the product of the common term times the algebraic sum of the unlike terms, plus the algebraic product of the unlike terms.

Examples

$$\begin{aligned}(a+3)\ (a+4) &= a^2 + a(3+4) + (3 \times 4)\\ &= a^2 + 7a + 12\end{aligned}$$

$$(a-5)\ (a-3) = a^2 - 8a + 15$$

$$(a-7)\ (a+3) = a^2 - 4a - 21$$

$$(x+7)\ (x-9) = x^2 - 2x - 63$$

A trinomial such as $x^2 + x\ (a + b) + ab$ is often expressed in the general form $x^2 + mx + n$ in which m represents the sum of the unlike terms and n, their product. *Factoring*, or finding the binomial factors, in trinomials of this form involves finding two quantities whose sum equals m and whose product equals n. After you determine these quantities, write the binomial factors by adding these quantities to the square root of the first term of the trinomial. Thus, for example, in $b^2 - 9b + 14$, the numbers -7 and -2 must be part of the binomial factors, for their sum is -9 and their product is $+14$. This means that the binomial factors are $(b - 7)$ and $(b - 2)$.

Examples

$$e^2 - 9e + 20 = (e-5)\ (e-4)$$

$$r^2 + 11r - 42 = (r+14)\ (r-3)$$

$$p^2 - p - 42 = (p-7)\ (p+6)$$

$$t^2 + 11t + 24 = (t+3)\ (t+8)$$

Square of a Binomial

When the product is of the form $(x + a)\ (x + a)$ or $(x + a)^2$, the middle term of the product is $2ax$, and the final term is a^2. Thus, $(x + a)^2 = x^2 + 2ax + a^2$.

Expressed in rule form, this relation is: The square of a binomial is the sum of the squares of the two terms added to twice their algebraic product.

Examples

$$(r+t)^2 = r^2+2rt+t^2$$
$$(k-2j)^2 = k^2-4jk+4j^2$$
$$(3e-d)^2 = 9e^2-6ed+d^2$$

Factoring polynomials of the type $x^2+2ax+a^2$, called perfect trinomial squares, depends largely on your ability to recognize them as such. A trinomial is a perfect square if two of its terms are perfect squares and the third term is twice the product of their square roots.

Examples

$$g^2+2gk+k^2 = (g+k)^2$$
$$p^2+2pq+q^2 = (p+q)^2$$
$$r^2-10r+25 = (r-5)^2$$

Other Binomial Products

When multiplying binomials of the type $mx+a$ and $nx+b$, whose product is $mnx^2+x(na+mb)+ab$, little time is saved by following a rule. However, knowing the rule for such multiplication may make it easier for you to factor the product of binomials of this type.

When multiplying any two binomials,

1. To find the first term of the product, multiply the first terms of the binomials by each other.
2. To find the second term of the product, obtain the product of the outer terms of the binomials and the product of the inner terms of the binomials and add them algebraically.
3. To find the third term of the product, find the algebraic product of the second terms of the binomials.

Thus, in finding the product of $(2e-3)(3e+4)$, you would proceed as follows:

First term: $2e \times 3e$, or $6e^2$.

Second term: $(2e\times 4)+(-3\times 3e)=8e-9e$ or $-e$.

Third term: -3×4, or -12.

The total product is therefore: $6e^2-e-12$.

Factoring trinomials of the form $mnx^2+(mb+na)x+ab$ usually involves a trial-and-error method. As you saw in the rule just given, the factors must be such that the product of the first terms of the binomials equals the first term of the trinomial, the product of the second terms of the binomials equals the third term of the trinomials, and the algebraic sum of the cross products is the second term of the trinomial. Therefore, the usual procedure is to try two binomials which satisfy the first and second conditions, then check to see if they also satisfy the third condition.

Example

Factor $6a^2-17a+12$.

Trial Factors		*Product*
$(3a-2)(2a-6)$	=	$6a^2-22a+12$ *Wrong* *middle term incorrect*
$(6a-1)(a-12)$	=	$6a^2-73a+12$ *middle term still incorrect*
$(6a-3)(a-4)$	=	$6a^2-27a+12$ *middle term still incorrect*
$(3a-4)(2a-3)$	=	$6a^2-17a+12$ *Right*

With practice, you can eliminate some elements in guessing the right combination. For example, you can tell by the signs of the trinomial whether the signs of the binomials will be both plus, both minus, or one plus and one minus. Do not use a binomial factor that contains a monomial factor. Remove monomial factors from the trinomial before extracting the binomial factors. This means that in the example just given it was not necessary to try the first combination $(3a-2)(2a-6)$, since 2 is a factor of $(2a-6)$ but not of the trinomial. Likewise, the third combination $(6a-3)(a-4)$ was unnecessary since 3 is a factor of $(6a-3)$, but not of the trinomial.

FRACTIONS

A fraction is an indicated division in which the *numerator* (number above the line) is the dividend and the *denominator* (number below the line) is the divisor. The ratio of two quantities (the value of the fraction) is unchanged if they are both multiplied or divided by the same number (not zero). Dividing both numerator and denominator by the same number is called reduction to lower terms.

Examples

1. $\frac{14}{21}=\frac{(2)(7)}{(3)(7)}=\frac{2}{3}$
2. $\frac{6ax}{9a}=\frac{(2x)(3a)}{(3a)(3a)}=\frac{2x}{3a}$
3. $\frac{a^2-b^2}{a^2-2ab+b^2}=\frac{(a+b)(a-b)}{(a-b)(a-b)}=\frac{a+b}{a-b}$

Note that a quantity can be used as a divisor only if it is a *factor* of the *complete* numerator and *complete* denominator.

4. $\frac{e^2-14e-51}{e^2-2e-15}=\frac{(e-17)(e+3)}{(e-5)(e+3)}=\frac{e-17}{e-5}$

Equivalent Fractions

To change a given fraction to an equivalent fraction with a new denominator, multiply both numerator and denominator by a number which will make the new denominator the desired value. For example, to change $\frac{1}{5}$ to 25ths, multiply both the denominator and numerator by 5. The equivalent fraction is $\frac{5}{25}$.

Examples

1. *Given:* $\frac{2a}{5b}=\frac{?}{10ab}$, *find the equivalent numerator.*

Dividing 10ab by 5b gives 2a, so both terms of the fraction must be multiplied by 2a to change to the desired denominator.

Thus, $\frac{2a}{5b}\times\frac{2a}{2a}=\frac{4a^2}{10ab}$ *and* $\frac{2a}{5b}=\frac{4a^2}{10ab}$

2. *Given:* $\frac{2m+1}{3m-5}=\frac{?}{6m^2-7m-5}$

The factors of $6m^2-7m-5$ *are* $(3m-5)(2m+1)$

Hence $\frac{(6m^2-7m-5)}{(3m-5)}=(2m+1)$

Multiplying,

$$\frac{2m+1}{3m-5}\times\frac{2m+1}{2m+1}=\frac{(2m+1)(2m+1)}{(3m-5)(2m+1)}=\frac{4m^2+4m+1}{6m^2-7m-5}$$

Addition and Subtraction of Fractions

Fractions may be combined (added or subtracted) only if they have the same common denominator. When you need to combine fractions with different denominators, first change them to equivalent fractions with the same denominator (called *common denominator*). The least common denominator (LCD) is best since the use of any other necessitates further reduction of the answer to lower terms. When dealing with polynomial denominators, factor each denominator as far as possible. The least common denominator is then the product of all the factors of the denominators, taking each factor the greatest number of times it occurs in any one denominator.

Examples

1. *Combine:* $\frac{x}{4}+\frac{3x}{5}-\frac{2x}{6}$

$L.C.D.=60$

$$\frac{(15)x+(12)3x-10(2x)}{60}$$

$$\frac{15x+36x-20x}{60}=\frac{31x}{60}$$

2. *Combine:* $\frac{3}{4xy}-\frac{2}{3x^2}$

$L.C.D.=12x^2y$

$$\frac{(3x)3-(4y)2}{12x^2y}=\frac{9x-8y}{12x^2y}$$

3. *Combine:* $\frac{1}{R^2-1}+\frac{3}{R^2-2R+1}$

$$=\frac{1}{(R+1)(R-1)}+\frac{3}{(R-1)(R-1)}$$

$$=\frac{1(R-1)+3(R+1)}{(R-1)(R-1)(R+1)}$$

$$=\frac{R-1+3R+3}{(R-1)(R-1)(R+1)}$$

$$=\frac{4R+2}{(R-1)^2(R+1)}$$

4. *Combine:* $3R+5-\frac{5}{2R+1}$

$$=\frac{3R+5}{1}-\frac{5}{2R+1}$$

$$=\frac{(3R+5)(2R+1)-5}{2R+1}$$

$$=\frac{6R^2+13R+5-5}{2R+1}=\frac{6R^2+13R}{2R+1}$$

The expression combined in example 4 is known as a *mixed expression* and the answer, as an *improper fraction*. Handle them as mixed numbers and improper fractions are handled in arithmetic.

Multiplication and Division of Fractions

To multiply two fractions together, find the product of the numerators and of the denominators. Frequently it is possible to reduce the result to lower terms by dividing out *factors* common to both numerator and denominator. This is sometimes called *cancellation of common factors.* To divide one fraction by another, invert the second fraction and multiply. Factoring of polynomial numerators and denominators is advisable in order to eliminate common factors from the result.

Examples

1. *Perform the indicated operations:* $\frac{4a}{3b}\times\frac{6ab}{2a^2}\div\frac{7ac}{9c^2}$

$$=\frac{4\cancel{a}}{\cancel{3b}}\times\frac{\overset{\cancel{2a}}{\cancel{6ab}}}{\underset{\cancel{a}}{\cancel{2a^2}}}\times\frac{\overset{9c}{\cancel{9c^2}}}{7a\cancel{c}}=\frac{36c}{7a}$$

Note the use of cancellation in this example. Wherever fractions are being multiplied, you can cancel out like terms. Note in the example that an a above the division line is cancelled and that the a^2 below is replaced by an a. Remember,

$$\frac{a}{a^2}=\frac{a}{aa}=\frac{a}{a}\times\frac{1}{a}=\frac{1}{a}$$

Therefore, when you see an a above and a^2 below, you can do the steps just outlined mentally, leaving, in this case, an a below the line. But since there is an a in the first fraction, this left-over a can be cancelled into it. Further cancellation removes practically two-thirds of the fraction.

2. Multiply: $\frac{m^2+5m+6}{m^2-4}\times\frac{m^2-6m+8}{m^2-9}$

$$=\frac{(m+2)\ (m+3)}{(m+2)\ (m-2)}\times\frac{(m-2)\ (m-4)}{(m+3)\ (m-3)}$$

$$=\frac{m-4}{m-3}$$

3. Perform the indicated operations:

$$\frac{5t^2+8t+3}{2t^2-6t}\times\frac{6t^2-18t}{5t^2-2t-3}\div\frac{t^2+5t+4}{t^2+3t-4}$$

$$=\frac{5t^2+8t+3}{2t^2-6t}\times\frac{6t^2-18t}{5t^2-2t-3}\times\frac{t^2+3t-4}{t^2+5t+4}$$

$$=\frac{(5t+3)\ (t+1)}{2t(t-3)}\times\frac{6t(t-3)}{(5t+3)\ (t-1)}\times\frac{(t-1)\ (t+4)}{(t+1)\ (t+4)}$$

$$=\frac{6t}{2t}=3$$

Complex fractions are fractions whose numerator or denominator, or both, contain fractions. They can be simplified by treating them as division problems.

Example

Simplify: $$\frac{R_1\left(\frac{R_2R_3}{R_2+R_3}\right)}{R_1+\frac{R_2R_3}{R_2+R_3}}$$

(The numbers 1, 2, 3 below and to the right of the R's are subscripts. They indicate that the R's are different quantities and are to be treated as if they were different letters.)

$$=R_1\left(\frac{R_2R_3}{R_2+R_3}\right)\div\left(R_1+\frac{R_2R_3}{R_2+R_3}\right)$$

$$=\frac{R_1R_2R_3}{R_2+R_3}\div\frac{R_1R_2+R_1R_3+R_2R_3}{R_2+R_3}$$

$$=\frac{R_1R_2R_3}{R_2+R_3}\times\frac{R_2+R_3}{R_1R_2+R_1T_3+R_2R_3}$$

$$=\frac{R_1R_2R_3}{R_1R_2+R_1R_3+R_2R_3}$$

FRACTIONAL EQUATIONS

The solution of fractional equations is not as difficult as you might think, for you can multiply both sides of an equation by the same number (axiom 2). Multiplying all the fractions by the common denominator, you obtain an equation involving whole numbers, which you can then solve as previously outlined.

Examples

1. Solve for a in the equation $\frac{6a}{7}-\frac{1}{3}=\frac{5a}{14}+\frac{13}{6}$

Given: $$\frac{6a}{7}-\frac{1}{3}=\frac{5a}{14}+\frac{13}{6}$$

Solution

42 is the common denominator

Multiply both sides by 42,

$$36a-14=15a+91$$

Add 14 to both sides,

$$36a=15a+105$$

Subtract 15a from both sides,

$$21a=105$$

Divide both sides by 21,

$$a=5$$

2. Solve for x in the equation, $\frac{8}{x}+\frac{1}{2}=2-\frac{7}{x}$

Given: $$\frac{8}{x}+\frac{1}{2}=2-\frac{7}{x}$$

Solution

Multiply both sides by 2x,

$$16+x=4x-14$$

Subtract 16 from both sides,

$$x=4x-30$$

Subtract 4x from both sides,

$$-3x=-30$$

Divide both sides by −3,

$$x=10$$

3. Solve for m in the equation $\frac{2m+1}{4m+1}=\frac{2m+5}{4m+7}$.

Given: $$\frac{2m+1}{4m+1}=\frac{2m+5}{4m+7}$$

Solution

Multiply both sides by (4m+1) (4m+7),

$$(2m+1)\ (4m+7)=(2m+5)\ (4m+1)$$

$$8m^2+18m+7=8m^2+22m+5$$

Subtract 8m² from both sides,

$$18m+7=22m+5$$

Subtract 7 from both sides,

$$18m = 22m - 2$$

Subtract 22m from both sides,

$$-4m = -2$$

Divide both sides by −4,

$$m = 0.5$$

4. *Solve for r_t in the equation,* $\frac{1}{r_t} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3}$

Given: $$\frac{1}{r_t} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3}$$

Solution

Multiply both sides by (r_t) (r_1) (r_2) (r_3),

$$r_1r_2r_3 = r_2r_3r_1 + r_1r_3r_t + r_1r_2r_t$$

Factor the right hand member,

$$r_1r_2r_3 = r_t(r_2r_3 + r_1r_3 + r_1r_2)$$

Divide both sides by $(r_2r_3 + r_1r_3 + r_1r_2)$,

$$\frac{r_1r_2r_3}{r_2r_3 + r_1r_3 + r_1r_2} = r_t$$

SIMULTANEOUS LINEAR EQUATIONS

As long as equations contain only one unknown, the simple rules for solution apply, but when you encounter a problem with two unknowns, you need to use *simultaneous linear equations*.

Simultaneous linear equations are two or more equations which contain only first powers of the unknown quantities and no products of unknowns. They are true for certain values of the unknowns. Among the various methods for finding the values of the unknowns are solution by graphing, elimination by addition or subtraction, elimination by substitution, and the use of determinants. Of these, the most common method is elimination by addition or subtraction. Therefore, it is well to understand this method.

If the letter you wish to eliminate has the same coefficient in the two equations, merely add (or subtract) the two equations. If the coefficients are different, you can make them the same by multiplying both members of each equation by some number, as you can see in the sample problem.

Example

Solve for e and i in the following pair of equations,

(1) $2e + 10i = 25$

(2) $5e - 8i = 46$

Solution

Multiply equation (1) by 4,

(3) $8e + 40i = 100$

Multiply equation (2) by 5,

(4) $25e - 40i = 230$

Add equation (4) to equation (3),

(5) $33e = 330$

Divide both sides by 33,

(6) $e = 10$

Substitute 10 for e in (1),

(7) $20 + 10i = 25$

Subtract 20 from both sides,

(8) $10i = 5$

Divide both sides by 10,

(9) $i = 0.5$

In this example, the coefficients of *i* were made equal by multiplying equation (1) by *4*, and equation (2) by *5*. You could have eliminated the *e's* by multiplying equation (1) by *5*, and equation (2) by *2*. While equations are not always in the form of (1) and (2), you can put them in such a form by simplifying and transposing.

If you have three equations with three unknown quantities, follow the same procedure of elimination. Eliminate one unknown by combining one pair of equations, then eliminate the same unknown from another pair of equations. This results in two equations with two unknowns, which you can solve as explained before.

Example

(1) $9i_1 - 4i_2 - 7i_3 = 4$

(2) $5i_1 + 3i_2 - 10i_3 = 3$

(3) $2i_1 - 5i_2 + 9i_3 = 2$

Solution

First, eliminate i_3 between (1) and (2)

Multiply equation (1) by 10,

(4) $90i_1 - 40i_2 - 70i_3 = 40$

Multiply equation (2) by 7,

(5) $35i_1 + 21i_2 - 70i_3 = 21$

Subtract equation (5) from equation (4),

(6) $55i_1 - 61i_2 = 19$

Next, eliminate i_3 from (1) and (3),

Multiply equation (1) by 9,

(7) $81i_1 - 36i_2 - 63i_3 = 36$

Multiply equation (3) by 7,

(8) $14i_1-35i_2+63i_3=14$

Add equation (7) and equation (8),

(9) $95i_1-71i_2=50$

Now you have two equations with two unknowns—

$55i_1-71i_2=50$ *equation (6)*
$95i_1-61i_2=19$ *equation (9)*

Multiply equation (6) by 19,

(10) $1045i_1-1159i_2=361$

Multiply equation (9) by 11,

(11) $1045i_1-781i_2=550$

Subtract equation (11) from equation (10),

(12) $-378i_2=-189$

Divide equation (12) by −378,

$$i_2=0.5$$

Substitute 0.5 for i_2 in equation (6),

$$55i_1-30.5=19$$

Add 30.5,

$$55i_1=49.5$$

Divide by 55,

$$i_1=0.9$$

Substitute 0.5 for i_2 and 0.9 for i_1 in equation (1),

$$8.1-2.0-7i_3=4$$

Collect terms,

$$6.1-7i_3=4$$

Subtract 6.1,

$$-7i_3=-2.1$$

Divide by −7,

$$i_3=0.3$$

Check by substituting in equations (1), (2), and (3),

$8.1-2.0-2.1=4$
$4=4$
$4.5+1.5-3.0=3$
$3=3$
$1.8-2.5+2.7=2$
$2=2$

FRACTIONAL EXPONENTS AND RADICALS

You have already noted that a number raised to the zero power is equal to one, and that a negative exponent is equal to the reciprocal of the quantity to the same positive exponent. Now consider the fractional exponent. Squaring the quantity $x^{1/2}$ gives $(x^{1/2})\ (x^{1/2})$. By adding the exponents $x^{(1/2+1/2)}$ as in any other multiplication with exponents, you find that the result is x^1, or x. By this reasoning, you can see that $x^{1/2}$ is the square root of x. In like manner, $x^{1/3}$ is the cube root of x, $x^{1/4}$ is $\sqrt[4]{x}$. The expression $x^{2/3}$ is read x to the two-thirds power. This means that x is raised to the power of the numerator of the fractional exponent, and reduced by the root of the denominator. Thus, $x^{2/3}$ is the cube root of x squared, which may be written $\sqrt[3]{x^2}$.

Simplifying Expressions with Radicals

Sometimes it is necessary to simplify an expression involving *radicals,* (square roots, cube roots, etc.) without changing its value. If you can, divide the quantity under the radical (called *radicand*) into two factors, and take the indicated root of one of the factors. This procedure is shown in the following examples.

Examples

1. $\sqrt{80}=\sqrt{16\cdot 5}=4\sqrt{5}$
2. $\sqrt{54}=\sqrt{9\cdot 6}=3\sqrt{6}$
3. $\sqrt[3]{16m}=\sqrt[3]{8\cdot 2m}=2\sqrt[3]{2m}$

When the quantity under the radical is a fraction, multiply both numerator and denominator by a quantity which will make it possible to take the indicated root of the denominator. This procedure is shown in the following examples.

Examples

1. $\sqrt{\frac{2}{3}}=\sqrt{\frac{2}{3}\cdot\frac{3}{3}}=\sqrt{\frac{6}{9}}=\sqrt{\frac{1}{9}\cdot 6}=\frac{1}{3}\sqrt{6}$
2. $\sqrt[3]{\frac{40}{x^2}}=\sqrt[3]{\frac{40}{x^2}\cdot\frac{x}{x}}=\sqrt[3]{\frac{40x}{x^3}}=\sqrt[3]{\frac{8}{x^3}\cdot 5x}=\frac{2}{x}\sqrt[3]{5x}$

Adding and Subtracting Radicals

You may add or subtract radicals directly if they have the same radicand and root and differ only in the coefficient. The examples which follow illustrate how to combine such radicals.

Examples

1. $4\sqrt{x}+3\sqrt{x}-2\sqrt{x}=5\sqrt{x}$
2. $\sqrt{4x}+5\sqrt{x}=2\sqrt{x}+5\sqrt{x}=7\sqrt{x}$
3. $\sqrt[3]{54}+\sqrt[3]{6\frac{3}{4}}=\sqrt[3]{27\cdot 2}+\sqrt[3]{\frac{27}{4}}=3\sqrt[3]{2}+\sqrt[3]{\frac{27}{4}\cdot\frac{2}{2}}$

$=3\sqrt[3]{2}+\sqrt[3]{\frac{54}{8}}=\sqrt[3]{2}+\sqrt[3]{\frac{27}{8}\cdot 2}$

$=3\sqrt[3]{2}+\frac{3}{2}\sqrt[3]{2}=4\frac{1}{2}\sqrt[3]{2}$

Multiplying Radicals

You can multiply two radicals together provided they have the same index (indicated root). You can do this by multiplying the coefficients together for the coefficient of the product and multiplying the radicands together for the radicand of the product. Study the following examples which illustrate the procedure used in multiplying radicals.

Examples

1. $(2\sqrt{3})\ (3\sqrt{2})=6\sqrt{6}$
2. $(7\sqrt{x})\ (2\sqrt{y})=14\sqrt{xy}$
3. $(2\sqrt{x})\ (3\sqrt{x})=6\sqrt{x^2}=6x$
4. $a\sqrt{b}\,(2\sqrt{3}+4)=2a\sqrt{3b}+4a\sqrt{b}$
5. $(5\sqrt{3}+2\sqrt{6})\ (5\sqrt{3}+4\sqrt{6})$

Multiply:

$$\begin{array}{r} 5\sqrt{3}+2\sqrt{6} \\ \underline{5\sqrt{3}+4\sqrt{6}} \\ 25\sqrt{9}+10\sqrt{18} \\ \underline{20\sqrt{18}+8\sqrt{36}} \\ 25\sqrt{9}+30\sqrt{18}+8\sqrt{36} \end{array}$$

$=25.3+30\sqrt{9\cdot 2}+8\cdot 6$

$=75+30\cdot 3\sqrt{2}+48$

$=123+90\sqrt{2}$

Dividing Radicals

To divide expressions containing radicals, use the opposite process — that is, divide coefficient by coefficient and radicand by radicand. Since this procedure often gives an expression with a radical in the denominator, a special process becomes necessary. This process, *rationalizing the denominator*, changes the denominator to a whole number, a fraction, or a mixed number.

You can rationalize the denominator by multiplying both the numerator and the denominator by a quantity chosen so that it will eliminate radicals from the denominator. In the case of monomials this is quite simple, and is the same process as discussed previously under *simplification of fractional radicands.*

Examples

$$\frac{\sqrt{5}}{\sqrt{6}}=\sqrt{\frac{5}{6}}=\sqrt{\frac{5}{6}\cdot\frac{6}{6}}=\sqrt{\frac{30}{36}}=\sqrt{\frac{1}{36}\cdot 30}=\frac{1}{6}\sqrt{30}$$

When the denominator is a binomial, rationalize by multiplying by the binomial with the sign between terms changed. This changed binomial is called the *conjugate* of the denominator. Since most applications of this process in radio and electrical circuits deal with binomials, examples involving more terms are not discussed.

Examples

1. $\dfrac{5}{3-\sqrt{2}}=\dfrac{5}{3-\sqrt{2}}\cdot\dfrac{3+\sqrt{2}}{3+\sqrt{2}}=\dfrac{15+5\sqrt{2}}{9-\sqrt{4}}=\dfrac{15+5\sqrt{2}}{9-2}$

   $=\dfrac{15+5\sqrt{2}}{7}$

   Here the denominator was $3-\sqrt{2}$ and the conjugate was $3+\sqrt{2}$.

2. $\dfrac{x}{\sqrt{x}+a}=\dfrac{x}{\sqrt{x}+a}\cdot\dfrac{\sqrt{x}-a}{\sqrt{x}-a}=\dfrac{x\sqrt{x}-ax}{\sqrt{x^2}-a^2}=\dfrac{x\sqrt{x}-ax}{x-a^2}$

3. $\dfrac{4-\sqrt{m}}{2+\sqrt{m}}=\dfrac{4-\sqrt{m}}{2+\sqrt{m}}\cdot\dfrac{2-\sqrt{m}}{2-\sqrt{m}}=\dfrac{8-6\sqrt{m}+\sqrt{m^2}}{4-\sqrt{m^2}}$

   $\dfrac{8+m-6\sqrt{m}}{4-m}$

4. $\dfrac{\sqrt{r}-\sqrt{s}}{\sqrt{r}+2\sqrt{s}}=\dfrac{\sqrt{r}-\sqrt{s}}{\sqrt{r}+2\sqrt{s}}\cdot\dfrac{\sqrt{r}-2\sqrt{s}}{\sqrt{r}-2\sqrt{s}}$

   $\dfrac{\sqrt{r^2}-3\sqrt{rs}+2\sqrt{s^2}}{\sqrt{r^2}-4\sqrt{s^2}}=\dfrac{r+2s-3\sqrt{rs}}{r-4s}$

Note that in each of the preceding examples the final result has no radicals in the denominator.

REAL AND IMAGINARY NUMBERS

Thus far you have dealt with roots of positive quantities. Consider now negative quantities appearing under radicals, such as $\sqrt{-x}$ or $\sqrt{-3}$. Obviously, there is no quantity which, when squared, will give $-x$ or -3. The indicated square roots of negative quantities are called *imaginary numbers.* In algebra, treat them as follows: $\sqrt{-x}=\sqrt{x}\cdot\sqrt{-1}=i\sqrt{x}$; $\sqrt{-3}=\sqrt{3}\cdot\sqrt{-1}=i\sqrt{3}$. In each case, i stands for the imaginary quantity, $\sqrt{-1}$. In electrical work the symbol j is used instead of i because the letter i is used to represent current. The quantity $\sqrt{-1}$ is usually referred to as the *j operator* and is used frequently in the solution of AC current problems.

Real and imaginary quantities can be graphically represented by four positions of a unit vector. Positive real numbers are plotted to the right of the origin, and negative real numbers to the left along the horizontal axis, which is known as the *axis of reals*. Imaginaries which have *positive* signs are plotted *above* the origin and negative imaginary quantities below the origin along the vertical axis which is known as the axis of imaginaries.

COMPLEX NUMBERS

A *complex number* is the sum or difference of a real quantity and an imaginary quantity. For example, $5-jb$ and $3+j2$ are complex numbers. The term complex number is really inappropriate because the system is not complex at all. *Rectangular notation* is a better designation. Two or more complex numbers can be added or subtracted by combining the real portions and imaginary portions separately as shown in the following examples.

Examples

Addition. (1) $\begin{array}{r} 5+j3 \\ \underline{2-j2} \\ 7+j \end{array}$ (2) $\begin{array}{r} 6-j4 \\ \underline{-5-j3} \\ 1-j7 \end{array}$

Subtraction. (1) $\begin{array}{r} 5+j3 \\ \underline{2-j2} \\ 3+j5 \end{array}$ (2) $\begin{array}{r} 6-j4 \\ \underline{-5-j3} \\ 11-j \end{array}$

To multiply complex numbers, use the same procedure as with binomials, except that, where j^2 occurs in the final result, you replace it by its equivalent, -1.

Examples

1. *Multiply $5+j3$ by $2-j2$*

$$\begin{array}{r} 5+j3 \\ \underline{2-j2} \\ 10+j6 \\ \underline{-j10-j^2 6} \\ 10-j4-j^2 6 \end{array}$$

$$=10-j4-(-1)(6)=10-j4+6=16-j4$$

You can divide one complex number by another by rationalizing the denominator, then dividing the real number into the numerator. You rationalize the denominator by multiplying both numerator and denominator by the *conjugate* of the denominator in the same manner as for radicals.

Examples

1. *Divide $(5+j3)$ by $(2-j2)$*

$$\frac{5+j3}{2-j2}\cdot\frac{2+j2}{2+j2}=\frac{10+j16+j^2 6}{4-j^2 4}=\frac{10+j16-6}{4+4}$$

$$=\frac{4+j16}{8}=\frac{1+j4}{2} \text{ or } 0.5\ +j2$$

2. *Divide $(6-j4)$ by $(-5-j3)$*

$$\frac{6-j4}{-5-j3}\cdot\frac{-5+j3}{-5+j3}=\frac{-30+j38-j^2 12}{25-j^2 9}$$

$$=\frac{-30+j38+12}{25+9}=\frac{-18+j38}{34}=\frac{-9+j19}{17}$$

QUADRATIC EQUATIONS

The degree of an equation in which the unknown has only positive integral exponents and does not appear in the denominator of a fraction is the same as its term of highest degree. The degree of a term in a letter means its exponent in that term. To illustrate, the degree of the term $4x^2y^3$ in x is the second. The equation $ax^2+bx+c=0$ is a second degree or quadratic equation in x. The following discusses the solution of equations of this form.

There are several methods you can use to solve a quadratic equation, such as graphing, completing the square, factoring, and using a formula derived from the general form $ax^2+bx+c=0$ by the completion of the square method. The formula is,

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

where a, b, and c are, respectively, the coefficient of x^2, the coefficient of x, and the term which does not contain x. Before applying the formula, always put the equation in the standard form $ax^2+bx+c=0$ by transposition. The following examples illustrate the processes for solving quadratic equations by the formula.

Examples

1: *Solve for x in the equation $3x^2+8x-10=0$.*

Solution

By the formula, $$x=\frac{-8\pm\sqrt{64+120}}{6}$$

$$=\frac{-8\pm\sqrt{184}}{6}=\frac{-8\pm 13.56}{6}$$

$$=\frac{-21.56}{6} \text{ or } \frac{5.56}{6}$$

$$=-3.59 \text{ or } 0.93$$

2. Solve for a in the equation $2a=3a^2-8$.

Solution

Subtract $3a^2$ *from both sides*

$$-3a^2+2a=-8$$

Add 8 to both sides

$$-3a^2+2a+8=0$$

By formula

$$a=\frac{-2\pm\sqrt{4+96}}{-6}=\frac{-2\pm\sqrt{100}}{-6}$$

$$=\frac{-2\pm10}{-6}$$

$$=\frac{-12}{-6}\text{ or }\frac{8}{-6}=2\text{ or }-1.33$$

3. Solve for R in the equation $\frac{R^2}{R+5}=\frac{R}{3}+\frac{5}{6}$

Solution

$$6R^2=2R(R+5)+5(R+5)$$

$$6R^2=2R^2+10R+5R+25$$

$$6R^2=2R^2+15R+25$$

$$4R^2-15R-25=0$$

$$R=\frac{15\pm\sqrt{225+400}}{8}$$

$$=\frac{15\pm\sqrt{625}}{8}=\frac{15\pm25}{8}$$

$$\frac{40}{8}\text{ or }\frac{-10}{8}=5\text{ or }-1.25$$

The quantity b^2-4ac which appears under the radical in the formula is called the *discriminant*. It indicates the type of roots. If b^2-4ac is positive, there are two real and unequal roots; if b^2-4ac is negative, the roots are imaginary and unequal; if b^2-4ac equals zero, the roots are real and equal.

LOGARITHMS

The *logarithm* of a quantity is the exponent, or the power, to which a given number, called the base, must be raised to equal that quantity. To illustrate, in the quantity $3^2=9$, the exponent, *2*, is called the logarithm of *9* to the base *3*. This relation is usually written $\log_3 9=2$. Any positive number greater than 1 may serve as a base. Two numbers have been selected, therefore, there are two systems of logarithms.

One system is the *natural* logarithm system whose base is the Greek letter epsilon (ϵ). The value of ϵ is approximately *2.718*. The other system is the *common* system of logarithms whose base is *10*. When no base is indicated for a logarithm, the base *10* is understood. Thus $\log 1000=3$ indicates that the base is *10*. For the natural system, the base ϵ may be written in or ln may be used as follows: $\log_\epsilon 5$ or ln *5*. Most of the work you do with logarithms will require the use of the common system of logarithms.

Common System of Logarithms

In the common system, logarithms that are exact powers of 10 are integers. Thus,

$\log 100=2$, since $10^2=100$

$\log 1000=3$, since $10^3=1000$

$\log 10000=4$, since $10^4=10000$

$\log 0.1=-1$, since $0.1=10^{-1}$

$\log 0.01=-2$, since $0.01=10^{-2}$

$\log 1=0$, since $0=10^0$

For numbers not exact powers of ten, the logarithm consists of two parts, an integral part (whole number) and a decimal part. The integral part is called the *characteristic* and the decimal part the *mantissa*. Thus, for example, $\log 595=2.7745$ (in words, the logarithm of *595* is *2.7745*). Here the characteristic is *2*, and the mantissa is *.7745*. The characteristic is found by inspection and the mantissa from tables of logarithms as explained in the following paragraphs.

DETERMINE THE CHARACTERISTIC. You can determine the characteristic by the following rules:

1. The characteristic of the logarithm of a number greater than 1 is positive and is one less than the number of digits to the left of the decimal point. For example, in the case of the log *595*, the characteristic is *2*, and for the log of *59.5*, the characteristic is *1*.

2. The characteristic of the logarithm of a number less than 1 is negative, and is equal to one more than the number of zeros immediately to the right of the decimal point. For example, for the log of *.0595*, the characteristic is *−2*, and for log *.00595*, the characteristic is *−3*.

When the characteristic is negative, do not put the minus sign in front of the logarithm, since the minus sign applies only to the characteristic and not to the entire logarithm. Instead, add *10* to the negative characteristic and subtract *10* at the end of the logarithm.

Thus the characteristic *−2* is written, *8.(man-*

tissa) -10, and the characteristic -3 is written, *7.*(*mantissa*) -10.

Another method of indicating that the characteristic is negative is to place the minus sign above the characteristic. With this method, the characteristics -2 and -3 are written $\overline{2}$.(*mantissa*) and $\overline{3}$.(*mantissa*).

FINDING THE MANTISSA. You can find the mantissa from the tables of logarithms on pages 549-550. Numbers which have the same figures in the same order and differ only in the position of the decimal point have the same mantissa in their logarithms. For example, the mantissa of *595* is *.7745*; the mantissa of *59.5* is also *.7745*.

After you determine by inspection the characteristic of the number as explained above, find the mantissa from the tables. The mantissa of the number is independent of the position of the decimal point, so you can disregard the decimal point in the number when finding the mantissa. The mantissas in the table are the decimal part of the logarithm and therefore should be preceded by the decimal point.

In four-place logarithm tables, the first column in the table contains the first two digits of the numbers whose mantissas are given in the table, and the top row contains the third digit. Thus, to find the mantissa of *595*, find *59* in the left-hand column and *5* at the top. In the column under *5*, and opposite *59*, is *7745*, the mantissa. The logarithm of *595* is then *2.7745*.

INTERPOLATION. To find the logarithm of a quantity with more than three digits, use the process called interpolation. Suppose you want to find the logarithm of *5956*. The tables do not give the mantissa for *5956*. However, they give the mantissas for *5950* and *5960*. (The mantissa for *5950* is the same as that for *595*. Likewise, the mantissa for *5960* is the same as that for *596*.) Since *5956* lies between *5950* and *5960*, its mantissa must lie between the mantissa for these two numbers.

By arranging the mantissas in the following tabular form,

Mantissa of *5960* = *.7752*
Mantissa for *5956* = *?*
Mantissa for *5950* = *.7745*

you can see that *5956* is *6/10* of the way between *5950* and *5960*. Obviously, the mantissa of *5956* must be *6/10* of the way between *.7745* and *.7752*. Since the difference between the two is *.0007* and since *6/10* of *.0007* is *.00042*, add *.00042* to *.7745* (the mantissa of *5950*). The result, *.77492*, is the mantissa of *5956*. Therefore, the logarithm of *5956* = *3.77492*.

ANTILOGARITHMS. The number corresponding to a given logarithm is called the antilogarithm of that number. It is written *antilog* or *log*$^{-1}$. To find the antilog, reverse the process for finding logarithms. The method is explained in the following examples.

Examples

1. To find the antilog (log^{-1}) of *1.8987*, first look in the logarithm table and locate the mantissa *.8987*. It is in line with the number *79* and under the column headed *2*. Thus the number corresponding with the mantissa *.8987* has the digits *792*.

To determine the location of the decimal point, reverse the rule for finding a characteristic. If the characteristic were zero, the decimal point would be placed after the first digit (*7.92*). Since the characteristic is *1*, count two places to the right and place the decimal point after the *9*. Thus the antilog of *1.8987* is *79.2*.

2. Find the antilog (log^{-1}) of *2.4325*.

The tables do not show the mantissa *.4325*; therefore, you must interpolate. The mantissa *.4325* lies between the two mantissas *.4330* and *.4314*. These mantissas correspond to *271* and *270* respectively. The difference between *.4330* and *.4314* is *.0016*, and the difference between *.4325* (the given mantissa) and the mantissa for *270* is *.0009*. Then find the number corresponding to *.4325* by adding $\frac{.0009}{.0016}$ of 1, or *.563* to *270*, thus giving the sum *270563*. Since the given characteristic is *2*, the antilog of *2.4325* is *270.563*.

Computations with Logarithms

MULTIPLICATION. To multiply two quantities, add their logarithms and find the antilog of the result as shown in the example which follows:

Example

Find the product of 6952 and 437.

Solution

$$\log (6952 \times 437) = \log 6952 + \log 437.$$

$$\log 6952 = 3.8421$$

$$\log 437 = 2.6405$$

Adding, $\log 6952 + \log 437 = 6.4826$

Find the antilog 6.4826.

The antilog $6.4826 = 3{,}038{,}000$.

Actual multiplication of 6952 by 437 would give the result 3,037,842. This indicates an error of 176 in over 3,000,000 or .006 of one percent. The error is due to the fact that the logarithm tables go to four places only. Greater accuracy would result in using five place tables. In general, though, accuracy obtained with four place tables is sufficient.

DIVISION. To divide two quantities, subtract the logarithm of the divisor from the logarithm of the dividend and find the antilog of the result as shown in the following example.

Example

Find the quotient of 6952 divided by 437.

Solution

$$\text{Log } (6952 \div 437) = \log 6952 - \log 437$$

$$\text{Log } 6952 = 3.8421$$

$$\text{Log } 437 = 2.6405$$

$$\text{Subtracting, } \log 6952 - \log 437 = 1.2016$$

Find antilog 1.2016

$$\text{Antilog } 1.2016 = 15.908$$

RAISING A QUANTITY TO A POWER. To raise a quantity to any power, multiply the logarithm of the quantity by the exponent, or the power, and find the antilog of the result as shown in the following examples.

Examples

1. *Find the value of* $(5.2)^6$

Solution

$$\text{Log } (5.2)^6 = 6 \times \log (5.2)$$

$$\text{Log } (5.2) = 0.716$$

$$\text{Log } (5.2)^6 = 6 \times (0.716) = 4.296$$

$$\text{Antilog } 4.296 = 19768$$

$$\text{Therefore, } (5.2)^6 = 19768$$

2. *Find the value of* $(3.7)^{1/5}$. *(Exponent is fractional)*

Solution

$$\text{Log } (3.7)^{1/5} = \frac{1}{5} \log (3.7)$$

$$\text{Log } (3.7) = 0.5682$$

$$\text{Then } \log (3.7)^{1/5} = \frac{1}{5} (0.5682) = 0.1136$$

$$\text{Antilog } 0.1136 = 1.2991$$

$$\text{Therefore, } (3.7)^{1/5} = 1.2991$$

3. *Find the value of* $(45.6)^{-3}$ *(Exponent is negative)*

Solution

$$\text{Log } (45.6)^{-3} = -3 \log (345.6)$$

$$\text{Log } 45.6 = 1.659$$

$$\text{Then, } \log (45.6)^{-3} = -3\ (1.659) = -4.977$$

Since logarithm tables list only positive values of mantissa, change -4.977 *to* $\overline{5}.023$ *(or* $5.023 - 10$*) by subtracting* -4.977 *from 10.*

$$\text{The antilog of } \overline{5}.023 = 0.0000010544.$$

$$\text{Therefore, } (45.6)^{-3} = .0000010544$$

FINDING THE ROOT OF A QUANTITY. To find the root of a quantity obtain the logarithm of the quantity, divide it by the index of the root, and find the antilog of the result as shown in the following example.

Example

Find the value of $\sqrt[3]{1.572}$

Solution

$$\text{Log } \sqrt[3]{1.572} = \frac{1}{3} \log 1.572$$

$$\text{Log } 1.572 = 0.19646$$

$$\text{Then } \log \sqrt[3]{1.572} = \frac{1}{3} (0.19646) = 0.06549$$

$$\text{Antilog } 0.06549 = 1.1611$$

$$\text{Therefore, } \sqrt[3]{1.572} = 1.1611$$

NATURAL LOGARITHMS

The natural system of logarithms uses a base ϵ which is approximately *2.718*. Natural logarithm tables give the complete logarithm rather than the mantissas only.

To find the natural log of a quantity which does not appear in the table and cannot be obtained by interpolation, or if there is no natural logarithm table, use the common logarithm table and the following relationships:

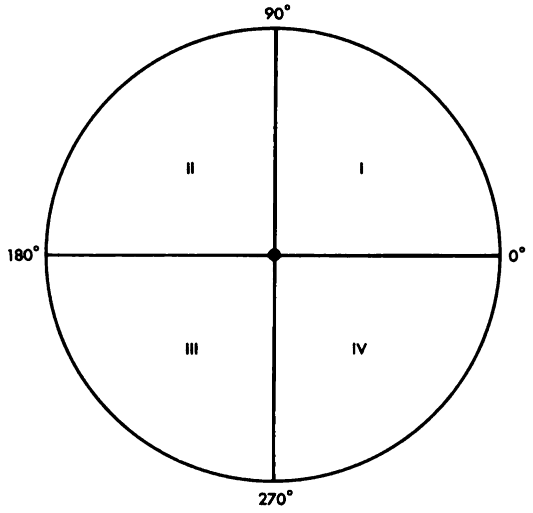
$$\log_e N = 2.3026 \log_{10} N$$

$$\log_{10} N = 0.4343 \log_e N$$

In multiplying, dividing, raising to powers, or finding roots with the natural system, use the same rules as in the common logarithm system.

TRIGONOMETRIC FUNCTIONS

Several special relationships, called *trigonometric functions*, hold true in a right triangle. Electrical problems when reduced to a right triangle can be easily and quickly solved by using tables based upon these functions. Before discussing the trigonometric functions, let us first see how angles are measured in mathematics and electronics.



Quadrant System

Angular Measurement

In navigation, you use a quadrant system with North as the reference point. The angles are measured in a clockwise direction through 360° A similar quadrant system is used in mathematics and electronics. However, a different reference point is used and the angles are generated in a counterclockwise direction.

As you can see in the illustration above, the portion of the horizontal axis to the right of the vertical axis is used as the reference line. Angles are generated about the origin in a counterclockwise direction from the reference line. Note that the quadrants are numbered in a counterclockwise direction also, as indicated in the accompanying illustration. The lines designated 0°, 90°, 180° and 270° are not in any particular quadrant, but are dividing lines between the quadrants.

Angles may also be measured in a clockwise direction using this quadrant system. An angle generated in a clockwise direction from the reference axis is called a negative angle. A negative angle can be generated through 360° also. In most cases, angles will fall in the first and fourth quadrants. In the illustration shown below, A is a positive angle of 30° and B is a negative angle of 30°.

Sine, Cosine, and Tangent

Let us now study the trigonometric functions-sine, cosine, and tangent. Suppose you draw a radius to some convenient point P on the circumference of a circle. As shown in the illustration at the top of the next page, P is in the first quadrant. From point P, drop a perpendicular to the horizontal axis, forming a right triangle. Label the radius Z, the vertical side X, the horizontal side R, and the angle between the radius and horizontal side θ. The radius of the circle is the hypotenuse of the right triangle and is often referred to as the radius vector.

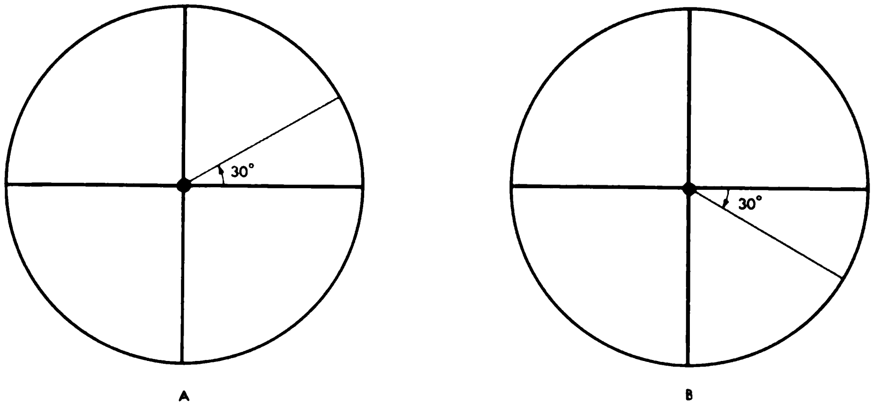
The *sine* (*sin*) *of an angle* in a right triangle is defined as the length of the side opposite that angle divided by the length of the hypotenuse. In the triangle in the illustration,

$$\sin\theta = \frac{X}{Z}$$

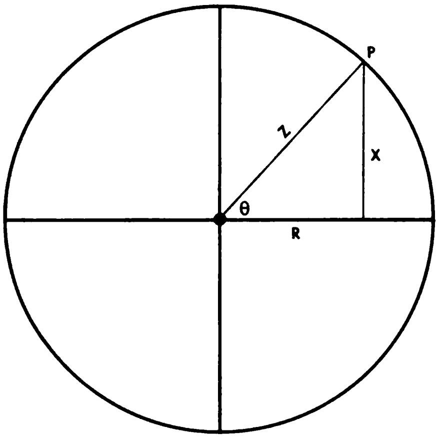
The *cosine* (*cos*) *of an angle* is defined as the length of the side adjacent to that angle divided by the length of the hypotenuse. So, in the triangle, you have

$$\cos\theta = \frac{R}{Z}$$

The *tangent* (*tan*) *of an angle* is defined as the length of the side opposite that angle divided by



Positive and Negative Angles



Finding Trigonometric Functions

the length of the side adjacent to that angle. So, in the triangle,

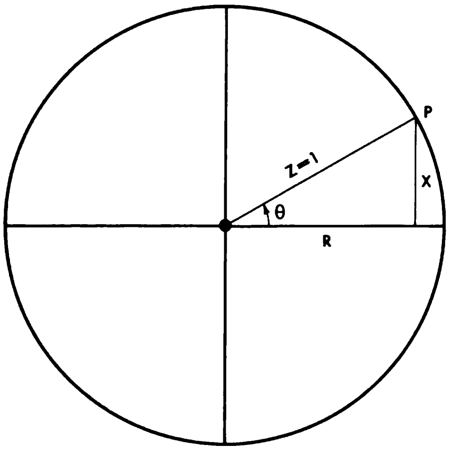
$$\tan \theta = \frac{X}{R}$$

Suppose you draw a circle with a radius one unit long. This is called a *unit circle* and is shown in the illustration below. In this circle,

$$\sin \theta = \frac{X}{Z} = \frac{X}{1} = X$$

Therefore, with the radius (or hypotenuse) equal to one, the length of X (the vertical component), will always be the sine of the angle.

Examine the illustration to the right. Note in A, where the angle *θ* is small, the vertical component X is also small. In B, the angle and the vertical component X have increased. In C, when the angle *θ* is 90°, the vertical component touches the circumference at point P. A perpendicular to the horizontal axis yields a vertical component equal in length to the radius.



Sin *θ* Equals Vertical Component in Unit Circle

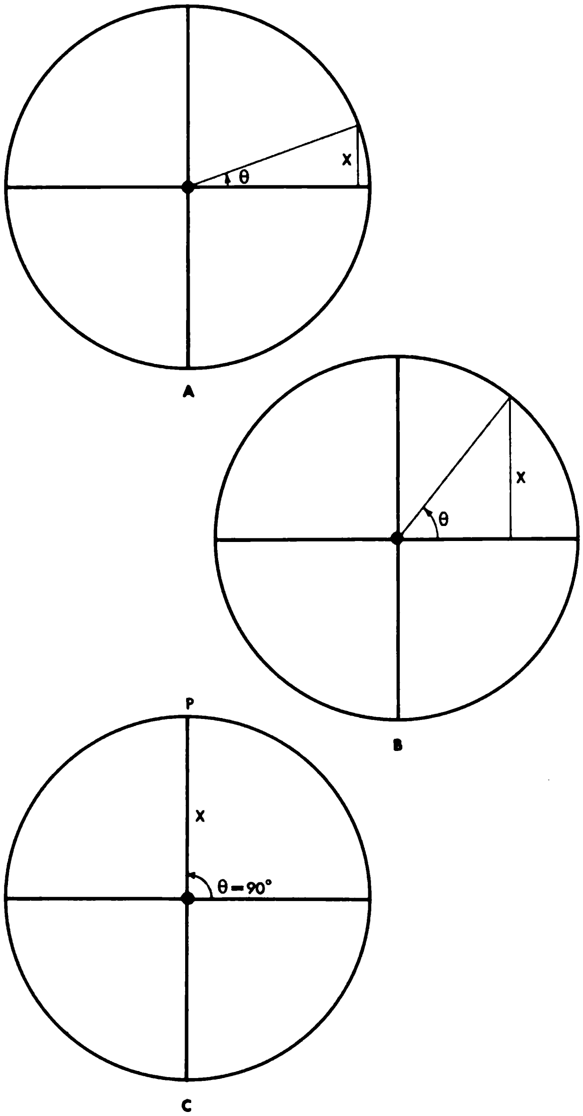
Therefore,

$$\sin 90° = \frac{X}{Z} = \frac{1}{1} = 1$$

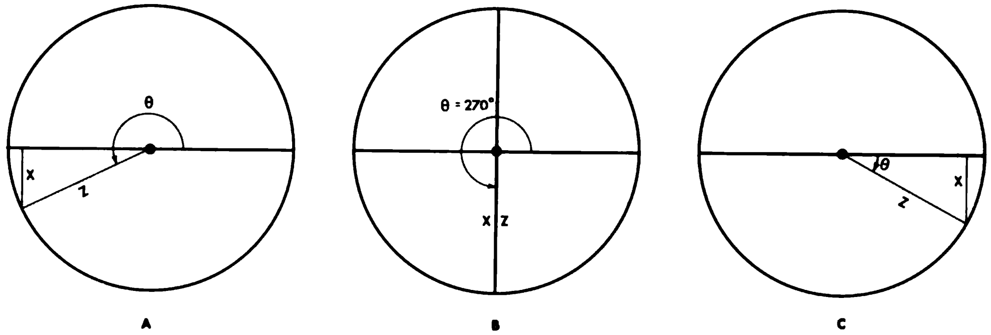
This is an important point to remember — that sin 90° is equal to 1. At 90°, you will get the maximum value of the sine — or the maximum value of the vertical component. Also remember that the sine of an angle will never be greater than 1, since the vertical component can never be greater than the radius.

As the angle *θ* increases from 90° to 180°, the vertical component decreases from 1 to 0 again. At 180°, the radius is lying upon the horizontal axis, and there is no vertical component.

$$\sin 180° = \frac{X}{Z} = \frac{0}{1} = 0$$



Effect of Angle on Vertical Component



Vertical Component of Angles in Quadrants III and IV

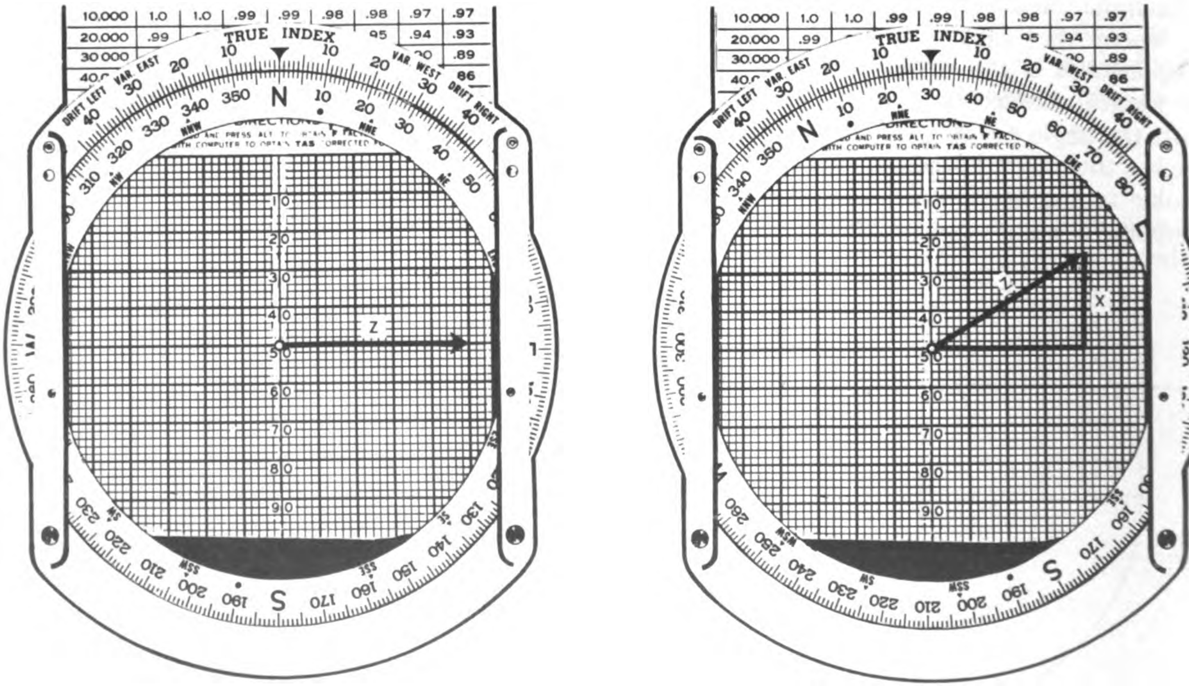
As θ increases from 180° to 270°, the perpendicular is drawn from the horizontal axis down to a point on the circumference. Note that when the right triangle is formed in the third quadrant, the vertical component lies below the horizontal axis, as shown at A of the illustration above. When the vertical component lies below the horizontal axis, the sine of the angle is negative. This conforms to the rules used in trigonometry, since the sines of angles from 180° to 360° are negative.

At B, with θ equal to 270°, note that the vertical component again is equal to the length of the radius Z. So sin 270° is also equal to one, but it has a negative value. Therefore,

$$\sin 270° = -1$$

Note also that a positive angle of 270° is equal to a negative angle of 90°. Any positive angle between 180° and 360° can also be referred to as a negative angle measured clockwise from the reference axis. In fact, this convention will be used in AC more frequently than measuring positive angles from 180° to 360°. Note at C that it is more convenient to refer to θ as a −30° angle rather than a positive angle of 330°. In the fourth quadrant, the sine of the angle will vary from −1 at 270° to 0 again at 360°.

Let us review briefly what happens to the value of the sine of an angle as the angle increases from 0° to 360°. The value of the sine increases from 0 at 0° to a maximum of 1 at 90°, and then decreases gradually to 0 at 180°. Its



A *Finding Sin 30° with DR Computer* **B**

value then increases in a negative direction until it reaches a maximum negative value at 270°. From 270° to 360° the negative value decreases until it finally falls to 0 again at 360°.

Using the DR Computer to Find the Sines of Angles

The DR computer can be used to find the approximate values of the sines of angles from 0° thru 360° by using the following procedure.

Place the "50" grid line under the grommet on the wind face of the computer and set the zero (N) marker under the index. Draw a vector from the grommet towards the right along the horizontal axis. Choose some convenient length on the grid and assign it a "unit" value (equal to 1). Choose as large a "unit" as possible, since the longer this line is, the more accurate your answers will be. (Five of the large squares will probably give the best results.) This unit will be the radius of the circle (Z). Each large square will represent 1/5 of the unit length or 0.2 as shown at A in the illustration on page 524.

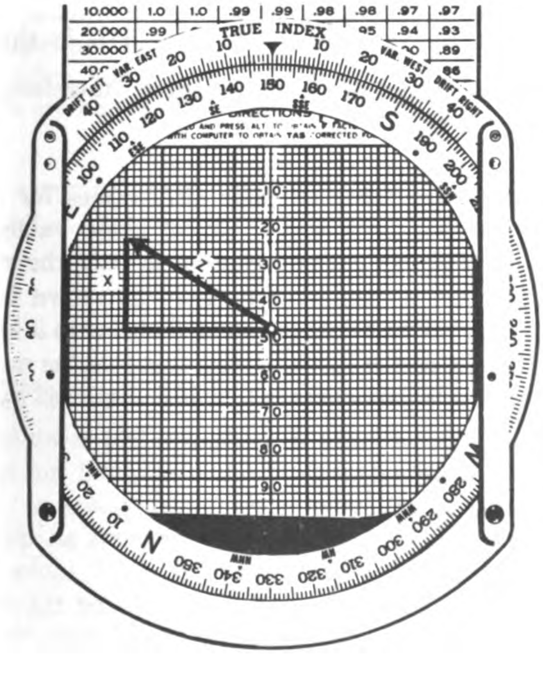
Now rotate the compass rose counterclockwise until 30° is under the index. Visualize a right triangle formed by the radius vector Z, the horizontal axis ("50" grid line) and a perpendicular dropped from the end of the radius vector (Z) to the horizontal axis. This perpendicular is the vertical component and represents the sine of 30°. The perpendicular distance is equal in length to 2½ squares, which is one-half of the 5 squares used to represent one "unit." Therefore, sin 30° = ½ = 0.5 as shown at B.

Putting 60° under the index, you have a new triangle formed and the vertical component represents approximately 0.87 of the unit length. Therefore sin 60° = 0.87. This process can be repeated for any angle you desire, and by measuring the length of the vertical component carefully, you will be able to solve for the sine of any angle with a high degree of accuracy.

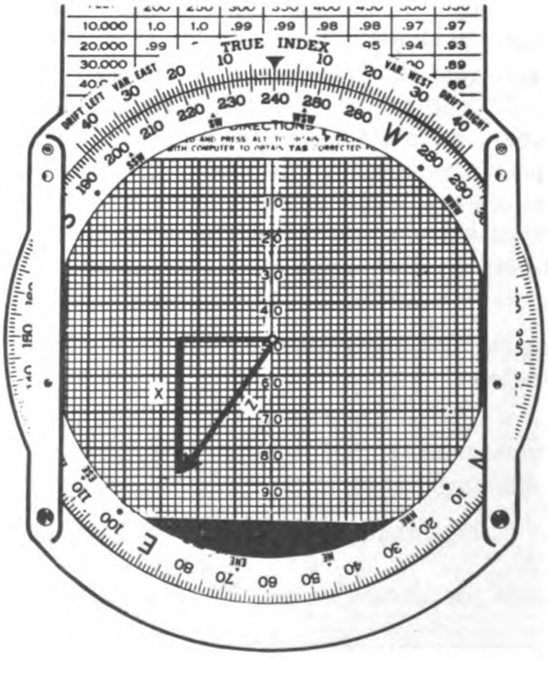
Note in the illustration to the left below that with 150° under the index, the radius vector falls in the second quadrant and the vertical component is again equal to 0.5. Therefore, sin 150° = 0.5.

With 240° set under the index, as shown in the illustration directly below, the radius vector falls in the third quadrant and the vertical component is 0.87 units long. However, since the vertical component is now below the horizontal axis, it is assigned a negative value. Therefore, sin 240° = −0.87.

With any angle from 180° to 360° set under the index, the vertical component falls under

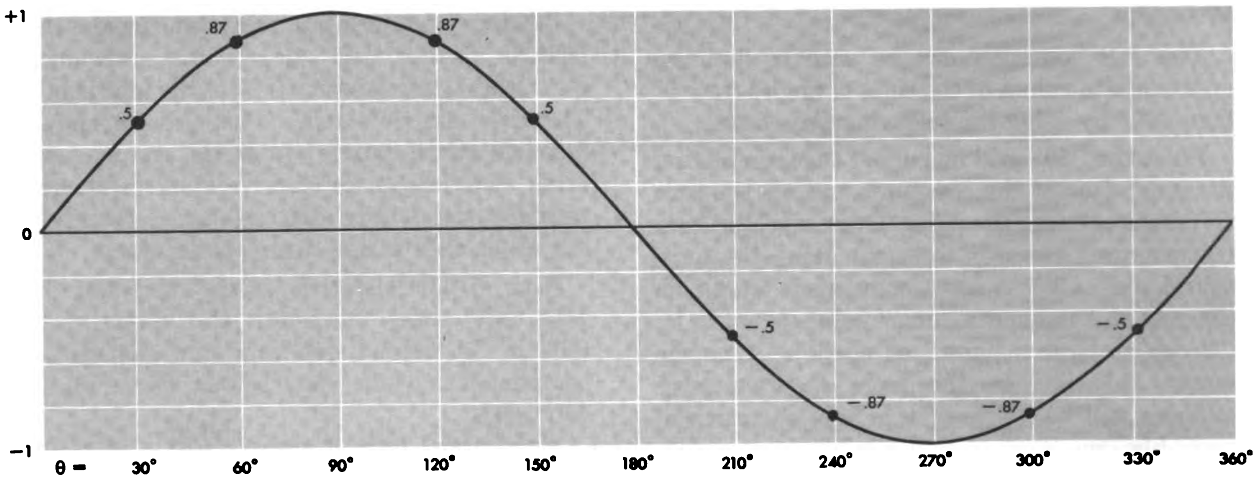


Finding Sin 150° with DR Computer



Finding Sin 240° with DR Computer

θ	0°	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
SIN θ	0	.5	.87	1	.87	.5	0	−.5	−.87	−1	−.87	−.5



Plotting the Values of Sin θ

the horizontal axis. So, all the angles from 180° to 360° have negative sines, according to the rules used in trigonometry.

The Sine Curve

In the table above, values for Sin θ for every 30° are shown. Plotting these values results in a curve as shown.

Notice that the horizontal axis is marked off every 30° and that its vertical axis can have a maximum value of plus one or minus one. When the points representing the sine of the various angles of θ are joined you have a mathematical curve known as a sine wave. This curve is used as a graphic representation of alternating currents and voltages used in electronics.

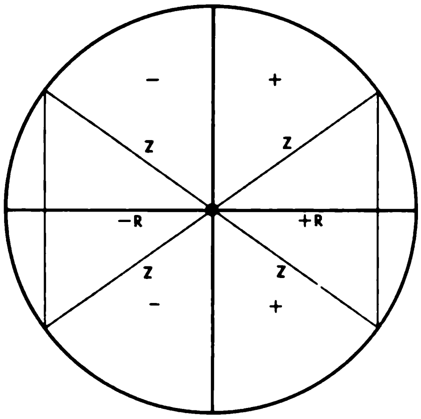
Using the DR Computer to Find the Cosines of Angles

Cos θ may be found on the DR computer in the same manner that has just been discussed for obtaining values of sin θ. In this case cos $\theta = \frac{R}{Z}$. Notice that when θ equals 0° the cos θ equals one, since $\frac{R}{Z} = \frac{1}{1} = 1$. When $\theta = 90°$, the cos θ is zero, since $\frac{R}{Z} = \frac{0}{1} = 0$.

As in the case of sin θ, the sign of cos θ depends on the quadrant in which the angle is. In the first and fourth quadrant the sign of cos θ is positive while in the second and third it is negative. Using the right triangle in the second quadrant, as shown in the illustration on the next page, you see that cos $\theta = \frac{R}{Z}$. Since it is in the second quadrant R is negative, therefore, cos $\theta = \frac{-R}{Z} = -\cos\theta$.

In the table on page 527, values for cos θ are shown for every 30°. If these values are plotted in the same manner as were the values for sin θ, you have a curve as shown in the illustration below the table. This curve is known as the cosine curve. Notice that it is the same as the sine curve except that it is displaced by 90°.

Although you can use the DR computer to find the values for sin θ, cos θ, and tan θ, you may sometimes have to find values more accurately. The computer gives only an approximation. You will find trigonometric tables from which the values of sin θ, cos θ, and tan θ may be obtained more accurately on pages 537-548 of the appendix.



Solution of the Right Triangle by the Pythagorean Theorem

Right triangles can be solved by using trigonometric functions. The lengths of the sides of the right triangle may also be found by means of the Pythagorean Theorem. This theorem expresses the relationship that exists between the lengths of the two sides and the hypotenuse of a right triangle no matter what the lengths of legs a and b or the size of angles A and B are. The Pythagorean Theorem states: *The square root of the sum of the squares of the legs of a right triangle equals the hypotenuse.* Expressed as an equation:

$$c = \sqrt{a^2+b^2}$$

The following example illustrates how the Phythagorean Theorem is used to find the hypotenuse when the value of the legs of a right triangle are known.

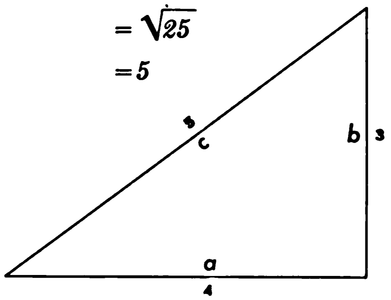
Example

If a = 4 inches and b = 3 inches, find the length of the hypotenuse c.

Solution:

Substituting in the equation:

$$\begin{aligned} c &= \sqrt{a^2+b^2} \\ &= \sqrt{4^2+3^2} \\ &= \sqrt{16+9} \\ &= \sqrt{25} \\ &= 5 \end{aligned}$$

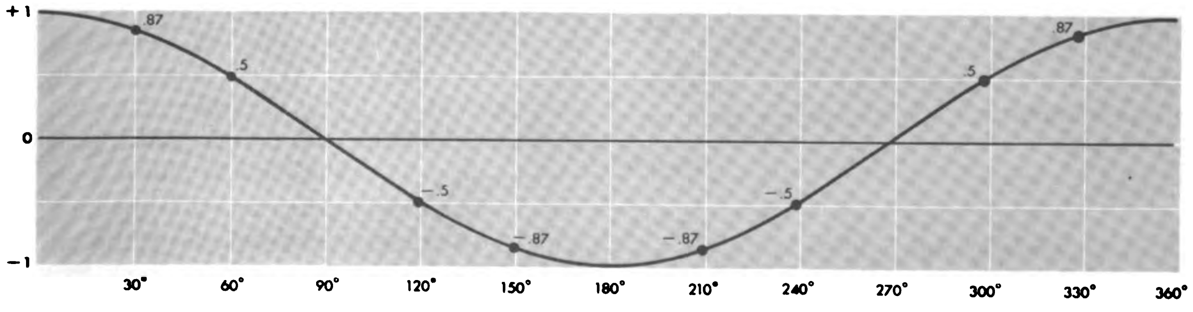


This solution to the right triangle is often used in AC for solving the impedance triangle.

Powers of Ten*

The powers of ten are sometimes termed the *engineer's shorthand.* A thorough knowledge of the powers of ten and the ability to apply the theory of exponents will greatly assist in determining an approximation. If a slide rule is used with the powers of ten, the average problem

θ	0	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
cos θ	1	.87	.5	0	−.5	−.87	−1	−.87	−.5	0	.5	.87



Plotting the Values of Cos θ

---

* Reprinted by permission from *BASIC MATHEMATICS FOR ELECTRONICS,* 2nd Ed., by Nelson M. Cooke. Copyright, 1960. McGraw-Hill Book Co., Inc.

reduces to the usual slide rule operations plus simple mental arithmetic. If a slide rule is not used for computation, the powers of ten enable one to work all problems by using convenient whole numbers. Either offers a convenient method for obtaining a final answer with the decimal point in its proper place.

Some of the multiples of 10 may be represented as follows:

Number	*Power of Ten*	*Expressed in English*
0.000001	$=10^{-6}$	= *ten to the negative sixth power*
0.00001	$=10^{-5}$	= *ten to the negative fifth power*
0.0001	$=10^{-4}$	= *ten to the negative fourth power*
0.001	$=10^{-3}$	= *ten to the negative third power*
0.01	$=10^{-2}$	= *ten to the negative second power*
0.1	$=10^{-1}$	= *ten to the negative first power*
1	$=10^{0}$	= *ten to the zero power*
10	$=10^{1}$	= *ten to the first power*
100	$=10^{2}$	= *ten to the second power*
1000	$=10^{3}$	= *ten to the third power*
$10,000$	$=10^{4}$	= *ten to the fourth power*
$100,000$	$=10^{5}$	= *ten to the fifth power*
$1,000,000$	$=10^{6}$	= *ten to the sixth power*

From the above, it is seen that any decimal may be expressed as a whole number times some negative power of ten. This may be expressed by the

RULE: To express a decimal as a whole number times a power of ten, move the decimal point to the *right* and count the number of places to the original point. The number of places counted is the proper negative power of ten.

Examples:

$$0.00687 = 6.87\times10^{-3}$$
$$0.0000482 = 4.82\times10^{-5}$$
$$0.346 = 34.6\times10^{-2}$$
$$0.08643 = 86.43\times10^{-3}$$

Also, it is seen that any large number may be expressed as some smaller number times the proper power of ten. This may be expressed by the

RULE: To express a large number as a smaller number times a power of ten, move the decimal point to the left and count the number of places to the original decimal point. The number of places counted will give the proper positive power of ten.

Examples:

$$435 = 4.35\times10^{2}$$
$$964,000 = 96.4\times10^{4}$$
$$6835.2 = 6.8352\times10^{3}$$
$$5723 = 5.723\times10^{3}$$

Multiplication with Powers of Ten

The law of exponents in multiplication, expressed in the general form, is

$$a^m \cdot a^n = a^{m+n} \text{ (where } a \pm 0)$$

This law is directly applicable to the powers of ten.

Example 1.

Multiply 1000 by 100,000.

Solution:

$$1000 = 10^3 \text{ and } 100,000 = 10^5$$
$$\text{then } 1000\times100,000 = 10^3\times10^5 = 10^{3+5} = 10^8$$

Example 2.

Multiply 0.000001 by 0.001.

Solution:

$$\begin{aligned} 0.000001 &= 10^{-6} \text{ and } 0.001 = 10^{-3} \\ \text{then } 0.000001\times0.001 &= 10^{-6}\times10^{-3} \\ &= 10^{-6+(-3)} \\ &= 10^{-6-3} \\ &= 10^{-9} \end{aligned}$$

Example 3.

Multiply 0.000037 by 600.

Solution:

$$\begin{aligned} 0.000037\times600 &= 3.7\times10^{-5}\times6\times10^{2} \\ &= 3.7\times6\times10^{-3} \\ &= 22.2\times10^{-3}, \text{ or } 0.0222 \end{aligned}$$

The student will find that by expressing all numbers between 1 and 10, times the proper power of ten, the determination of the proper place for the decimal point will become a matter of inspection.

Division with Powers of Ten

The law of exponents in division may be summed up in the following general form:

$$\frac{a^m}{a^n} = a^{m-n} \text{ (where } a \pm 0)$$

Example 1.

$$\frac{10^5}{10^3} = 10^{5-3} = 10^2$$

$$\text{or} \quad \frac{10^5}{10^3} = 10^5\times10^{-3} = 10^2$$

Example 2.

$$\frac{72,000}{0.0008} = \frac{72\times10^3}{8\times10^{-4}} = \frac{72}{8}\times10^{3+4} = 9\times10^7$$

$$\text{or} \quad \frac{72,000}{0.0008} = \frac{72\times10^3}{8\times10^{-4}} = \frac{72}{8}\times10^3\times10^4 = 9\times10^7$$

Example 3.

$$\frac{169\times10^5}{13\times10^5}=\frac{169}{13}\times10^{5-5}=13\times10^0=13\times1=13$$

$$\text{or}\quad\frac{169\times10^5}{13\times10^5}=13$$

It is apparent that like exponents in numerator and denominator may be canceled. Also, the student will note that powers of ten may be transferred at will from denominator to numerator, or vice versa, if the sign of the exponent is changed when the transfer is made.

Combined Multiplication and Division

This is most conveniently accomplished by alternately multiplying and dividing until the problem is completed.

Example.

$$\text{Simplify}\ \frac{0.000644\times96{,}000\times3300}{161{,}000\times0.00000120}$$

Solution. First convert all numbers in the problem to numbers between 1 and 10, times their proper power of 10, thus:

$$\frac{6.44\times10^{-4}\times9.6\times10^4\times3.3\times10^3}{1.61\times10^5\times1.2\times10^{-6}}=\frac{6.44\times9.6\times3.3\times10^4}{1.61\times1.2}$$

The problem as now written consists of multiplication and division of simple numbers. If the remainder of the problem is solved by slide rule, rough multiplication and division may be carried along mentally with no danger of misplacing the decimal point. If the problem is solved without the aid of a slide rule, there are no small decimals and no cumbersome large numbers to handle.

Instead of first finding the product of the numerator and dividing it by the product of the denominator, it is best to divide and multiply alternately. Thus, we divide 6.44 by 1.61 to obtain 4. Then we multiply this 4 by 9.6 to obtain 38.4. The 38.4 is then divided by 1.2 which results in a quotient of 32. Finally, 32 is multiplied by 3.3 which results in a product of 105.6. Because we still have a factor of 10^4, the answer is $105.6\text{x}10^4$. If we desire to express the answer in powers of ten, it would be written $1.056\text{x}10^6$; but written out, without the power of ten, it would be 1,056,000.

The method of alternately dividing and multiplying offers the slide rule operator the advantage of working the problem straight through without the necessity of jotting down the product of the factors of the numerator before proceeding to find the product of the denominator factors.

Reciprocals

In radio and electrical problems, many formulas are used that involve reciprocals, such as $\frac{1}{R_t}=\frac{1}{R_1}+\frac{1}{R_2}$, $X_c=\frac{1}{2\pi fC}$, $F=\frac{1}{2\pi\sqrt{LC}}$, etc.

The reciprocal of a number is 1 divided by that number. Such problems present no difficulty if the powers of ten are used properly.

Example 1.

$$\text{Simplify}\ \frac{1}{40{,}000\times0.00025\times125\times10^{-6}}$$

Solution: First convert all numbers in the denominator to numbers between 1 and 10, times their proper power of ten, thus:

$$\frac{1}{4\times10^4\times2.5\times10^{-4}\times1.25\times10^{-4}}=\frac{10^4}{4\times2.5\times1.25}$$

Multiplying the factors of the denominator results in $\frac{10^4}{12.5}$

Instead of writing out the numerator as 10,000 and then dividing by 12.5, the numerator is written as two factors in order to better divide mentally. The problem may be written:

$$\frac{10^2\times10^2}{12.5}\ \text{or}\ \frac{100}{12.5}\times10^2=8\times10^2$$

This method is of particular advantage to the slide rule operator because of the ease of estimating the number of figures in the final result.

If the final result is a decimal, rewriting the numerator into two factors allows fixing the decimal point with the least effort.

Example 2.

$$\text{Simplify}\ \frac{1}{625\times10^4\times2000\times64{,}000}$$

Solution: First convert all numbers in the denominator to numbers between 1 and 10, times their proper power of ten, thus:

$$\frac{1}{6.25\times10^6\times2\times10^3\times6.4\times10^4}=\frac{10^{-13}}{6.25\times2\times6.4}$$

Multiplying the factors in the denominator results in $\frac{10^{-13}}{80}$

Instead of writing out the numerator as 0.0000000000001 and dividing it by 80, the numerator is written as two factors in order to better divide mentally. The problem may be written as follows:

$$\frac{10^2\times10^{-15}}{80}, \text{ or } \frac{100}{80}\times10^{-15}=1.25\times10^{-15}$$

If the value of the denominator product was over 100 and less than 1000, we should break up the numerator so that one of the factors would be 10^3 or 1000, and so on. This method will always result in a final quotient of a number between 1 and 10, times the proper power of ten.

Metric Prefixes

Metric prefixes are abbreviations for very large or very small numbers. The metric prefixes most common in electronics are:

Prefix	*Symbol*	*Value*		
Kilo	*K*	*1000*	*or*	10^3
Mega	*M*	*1000000*	*or*	10^6
Milli	*m*	$\frac{1}{1000}$	*or*	10^{-3}
Micro	*μ*	$\frac{1}{1000000}$	*or*	10^{-6}
Micro-micro	*μμ*	$\frac{1}{1000000000000}$	*or*	10^{-12}

To illustrate the use of metric prefixes, consider the following examples:

Example 1.

1,000,000 ohms can be written as $1\times10^6\Omega$, but since the symbol for 10^6 is "M," we may write 1 MΩ (pronounced "1 Megohm").

Example 2.

.002 amperes can be written as 2×10^{-3}, but since the symbol for 10^{-3} is "m," we may write 2 ma (pronounced "2 milliamps").

Example 3.

Change 260 milliamperes to amperes:

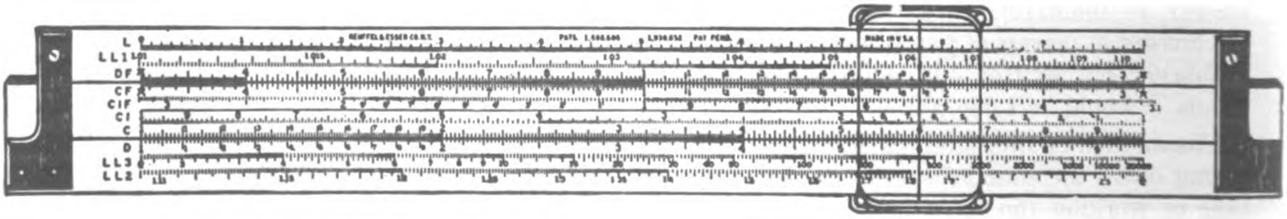
260 milliamperes can be expressed as 260×10^{-3} amp or, by using the rules discussed under Powers of Ten, .260a. Note that in this case, the movement of the decimal point three places to the left has accounted for the 10^{-3}

USE OF THE SLIDE RULE

Paper and pencil methods of making computations are at best slow and often introduce unintentional errors. Mathematical computations can be done very quickly and quite accurately by the use of a slide rule. This instrument consists of two rulers, one sliding back and forth on the other. It can be used in solving many types of problems. It is only a tool, however, and without you, it can do nothing.

The log-log decitrig slide rule shown below can indicate the final answer of a computation in the form of three or four digits. There are many rules for placing the decimal point. However, the surest method is to know the approximate answer, and from it locate the decimal point in the exact answer provided by the slide rule. For example, in multiplying 513 by 13, the sequence of digits is too confusing for mental calculation. According to the slide rule, the digits in the answer are 667. To locate the decimal, round off the numbers 513 and 13 to 500 and 10, respectively, and mentally calculate the product. From this product, 5000, you can determine that the decimal point should be located so that the answer given by slide rule reads 6670. By adopting such a rule for locating the decimal, you can avoid answers that are unreasonable.

The slide rule consists of several logarithmic scales so arranged that you can make computations by setting a movable portion, the *slide*, in a certain position relative to the fixed part, the *body*, and reading the answer. The types of computations that are possible with a slide rule depend on the type of slide rule. Usually they include multiplication, division, squaring, cubing, extracting square and cube roots, find-



Log-Log Decitrig Slide Rule

ing logarithms, and finding trigonometric functions. These processes are discussed briefly here, but before taking up the processes, it is well to know the meaning of several terms. The slide and body have already been identified. The glass runner is called the *indicator* and the mark upon it is the *hairline*. The mark associated with the number one on the body or slide is the *index*. Two positions on different scales are opposite each other when the hairline covers both positions at the same time.

Multiplication

Use the C and D scales to find the product of two numbers. To perform this operation, set the index of the C scale opposite the multiplicand on the D scale, and read the product on the D scale, opposite the multiplier on the C scale. This gives a number of three or four digits. Place the decimal point in accordance with a rough estimate. Note that the C scale has two indexes, one at each end. If you find that, when you set the left-hand index opposite the multiplicand, the multiplier on the C scale is beyond the index of the D scale, use the right-hand index of the C scale and proceed as explained.

Division

In dividing two numbers, set the slide so that the divisor on the C scale is opposite the dividend on the D scale. The number on the D scale opposite the index of the C scale is then the quotient.

Squaring and Extracting Square Root

To square a number, set the indicator so that the hairline covers the number to be squared on the D scale, then read the result as the opposite number on the A scale. (C and B scales may be used in like manner.)

To extract the square root of a number, set the hairline over the number to be squared on the A scale and read the root opposite it on the D scale. Since the A scale is a double scale, the same number appears on it at two places. For accurate results, use the *left-hand* portion of A for numbers between 1 and 10, and the *right-hand* portion for numbers between 10 and 100. In general, use the *left-hand* portion if the number has an *odd* number of digits and the *right-hand* portion if the number has an *even* number of digits to the left of the decimal point.

Logarithms

The L scale gives the mantissa of the common logarithm for the opposite number on the D scale. Determine the characteristic in the usual manner from the number of digits.

Find the antilog by locating the number on the D scale opposite the mantissa on the L scale and placing the decimal point as determined from the characteristic.

Reciprocals

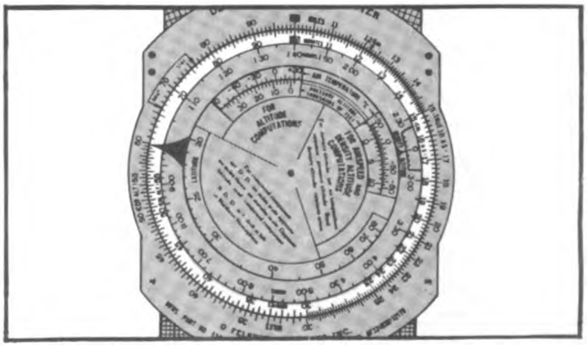
The CI scale is made up of the reciprocals of the opposite numbers on the C scale. The D and DI scales give the same information.

Trigonometric Functions

SIN. To find the sine of an angle, set the hairline to the black number (representing the size of the angle in degrees) on the S scale and read the sine from the C scale. Use the ST scale if the angle is less than any angle shown on the S scale. When you use the ST scale, the left index of C is 0.01 and the right index is 0.1; when you use the S scale, the left index is 0.1 and the right index is 1.0.

COS. You can find the cosine from the same scales. Since cos $\theta = \sin(90° - \theta)$, it is possible to find the cosine of an angle by finding the sine of 90° minus that angle. The *red* figures on the S scale are the values of $\sin(90 - \theta)$ or cos θ. Thus to find cos θ, set the hairline to the red number on S scale representing θ and read the opposite number on the C scale.

TAN. Use the following method in finding the tangent of an angle. For small angles, set the hairline to the angle on the ST scale and read the tangent on the C scale. The values on the C scale range from 0.01 at the left index to 0.1



The Fixed or Outer Scale Is Called the "Miles" Scale at the right. For angles from 5.74° to 45°, set the hairline to the black figure on the T scale, and read the tangent on the C scale. These values range from 0.1 to 1.0. For angles greater than 45°, use the red figures on the T scale, and read the result from the CI scale. These values range from 1 to 10.

Many other computations are possible with the slide rule. However, those given cover the bulk of calculations you will have to perform with a slide rule.

THE DR COMPUTER AS A SLIDE RULE

The face of the DR computer is nothing more than a circular slide rule. You can, therefore, use it to solve problems involving multiplication and division.

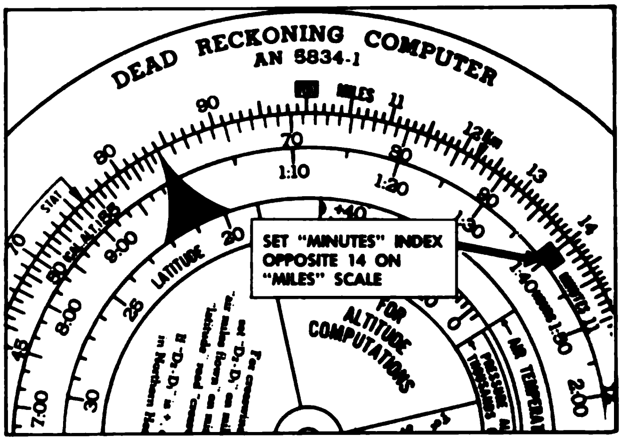
The Scales

The face of the DR computer is like that of a slide rule. The miles scale, as shown in the illustration on the previous page, is similar to the D scale of the slide rule. The 10 is considered to be the left index of this scale. On the slide rule the left index on the D scale is 1.

The rotating inner scale (minutes scale) performs the same function as the slide of the slide rule does. The 10, as shown in the illustration below, is also the index of this scale. The divisions on both scales are exactly the same as on the slide rule and are therefore read in the same manner.

Multiplication

You can multiply two numbers together in the same manner as with a slide rule. For example, to multiply 14 and 46 proceed as follows:


DEAD RECKONING COMPUTER
AN 5834-1
SET "MINUTES" INDEX OPPOSITE 14 ON "MILES" SCALE
FOR ALTITUDE COMPUTATIONS

To Multiply Two Numbers Such as 14 and 46

1. Set the index of the minutes scale under 14 on the fixed scale as shown in the illustration above.

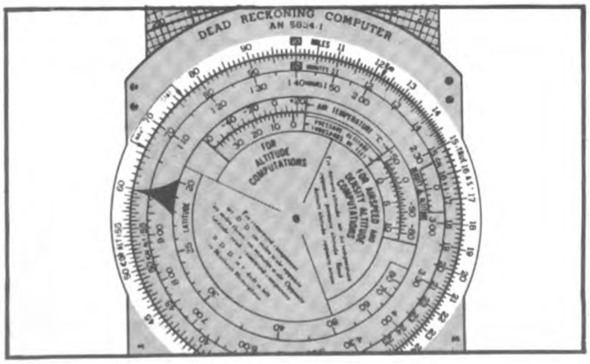
2. Opposite 46 on the minutes scale read the answer, 644, on the miles scale as shown in the illustration below.

Division

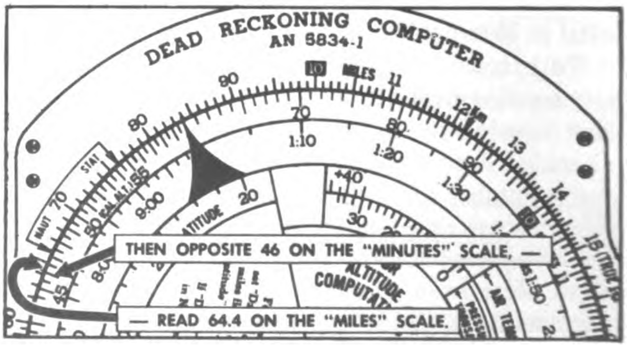
Division is the reverse of multiplication. Therefore, to divide 180 by 15, proceed a follows:

1. Under the dividend on the miles scale (180) set the divisor on the minutes scale (15), as shown in the illustration to the right above.

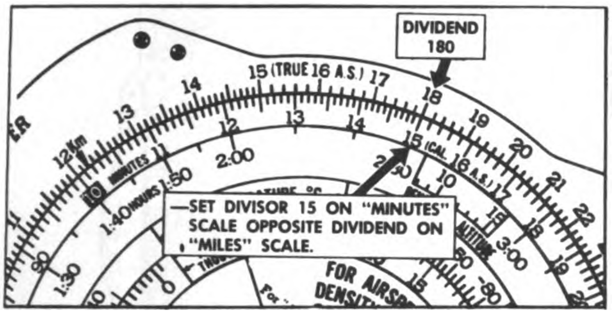
2. Opposite the index, 10, on the minutes scale read the quotient, 12, on the miles scale. This result is also shown in the illustration to the right.



The Rotating or Inner Scale Is Called the "Minutes" Scale


DEAD RECKONING COMPUTER
AN 5834-1
THEN OPPOSITE 46 ON THE "MINUTES" SCALE, —
— READ 64.4 ON THE "MILES" SCALE.

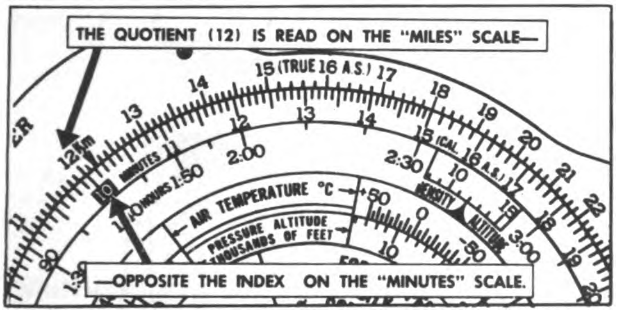
Since 14 x 46 Is Approximately 15 x 40 Which Equals 600, Obviously This Should Be 644, Hence 14×46=644



DIVIDEND 180

—SET DIVISOR 15 ON "MINUTES" SCALE OPPOSITE DIVIDEND ON "MILES" SCALE.

To Divide 180 by 15



THE QUOTIENT (12) IS READ ON THE "MILES" SCALE—

—OPPOSITE THE INDEX ON THE "MINUTES" SCALE.

VECTOR CONCEPTS

Vector quantities may be represented graphically by a *directed line.* The length of the line is scaled so that its magnitude corresponds to the magnitude of the vector quantity. The direction of the line is the same as the direction of the vector quantity. The convention used to indicate the direction of the line, or vector, is as follows:

1. All angles are measured from the horizontal axis which lies to the right of the vertical axis.

2. All angles are considered *positive* if they are measured in a *counterclockwise* direction from the horizontal axis.

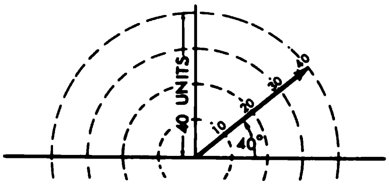
3. All angles are considered *negative* if they are measured in a *clockwise* direction from the horizontal axis.

A vector quantity whose magnitude is 40 units in a direction of 40° is shown graphically at the top right of this page. Note that the arrowhead denotes the direction of the vector.

Each particular magnitude and direction describes one, and only one vector. Vectors expressed in this fashion, that is, vectors expressed by merely stating the magnitude and direction, are said to be expressed in *polar form.* The vector shown in the illustration can be expressed conveniently in polar form as follows:

$$40\underline{/40^\circ}.$$

You read this vector as 40 at an angle of 40°.



There are two additional points concerning vectors that are worth considering. One is that a vector may be moved to a new position which is parallel to its original position without changing the vector. The second is that a vector may be moved to a new position anywhere along its axis without altering the vector. These changes, you will note, do not alter the properties of the vector (direction and magnitude), therefore, they do not change the vector. Such changes in position may be made, however, for purposes of calculation.

Adding Vectors

Vectors are added by plotting one vector, then placing the tail of the second vector at the head of the first and going in the direction indicated. The sum or resultant of adding vectors is always measured from the origin.

Vectors in polar form cannot be added algebraically, but can be added on the DR computer. For example, to add the vectors $10/0^\circ + 20/90^\circ$, proceed as follows:

1. Set the grommet of the DR computer on the top line of the grid. Then set 0° (the first direction) under the true index. Straight down from the grommet mark off ten units (the first magnitude) as shown in the first illustration on the next page.

2. Rotate the compass rose until 90° (the second direction) is under the true index. Straight down from the first plotted point, mark off 20 units (the second magnitude). You may have to move the grid until the first plotted point is on the horizontal axis so that a more accurate 20 units can be measured straight down from the first plotted point, as shown in the second illustration on the next page.

3. Rotate the compass rose until the second plotted point is over the vertical axis and be

sure that the grommet is back in its original position on the horizontal axis. In the bottom illustration on this page, read 21 units between the grommet and the second plotted point below it on the vertical axis. Then read the resulting direction under the true index (63°). The sum of the two vectors may therefore be written $21\underline{/63^\circ}$.

If one of the directions to be added is a negative angle, such as -20°, this angle would have to be subtracted from 360° in order that the positive equivalent of -20° could be put under true index.

Subtracting Vectors

If Z represents a positive vector, 40 units in length and at an angle of 50°, the negative counterpart of this vector is one of the same magnitude but with a reciprocal direction. Thus, if

$$Z=40\ \underline{/50^\circ}$$

then

$$-Z=40\ \underline{/50^\circ}+180^\circ$$

$$-Z=40\ \underline{/230^\circ}$$

The vectors Z and $-Z$ are shown in the illustration on top of the next page.

Occasionally, you will need to subtract two vectors in order to obtain the desired solution of a problem. As an example, let

$$Z=R+X_L$$

where $Z=55\underline{/58^\circ}$ and $R=30\underline{/0^\circ}$ and you wish to find X_L. You can transpose the R term in the given equation obtaining

$$X_L=Z-R.$$

Since Z and R are vector quantities, you must perform a vector subtraction of these two quantities to solve for X_L. This can be done in the following manner:

Change the equation to read the *sum* of the positive vector and the negative of the vector to be subtracted.

$$X_L=Z+(-R)$$

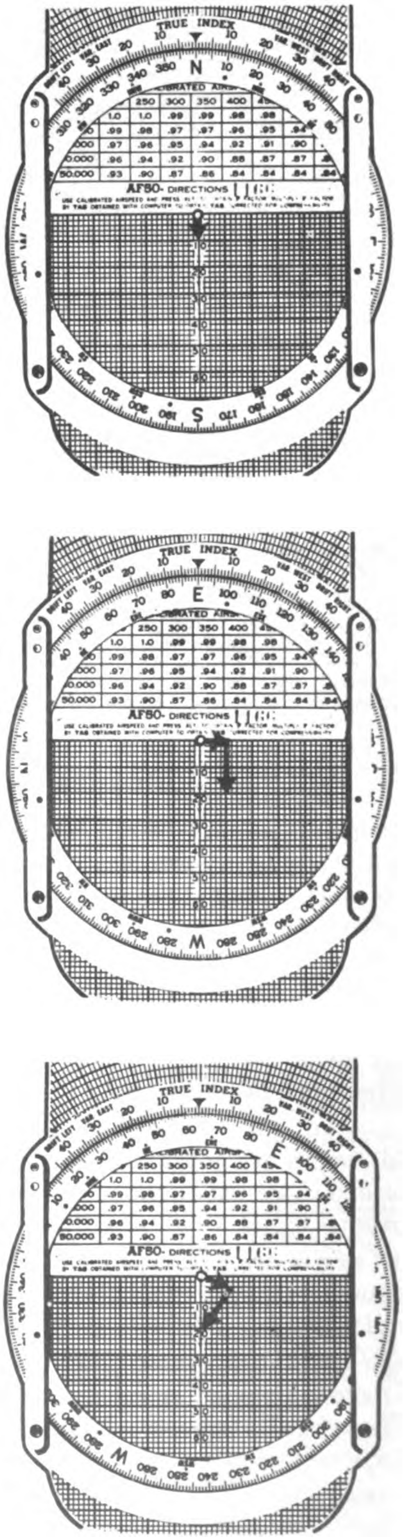
Since $\quad -R=+(-R)$

In the example,

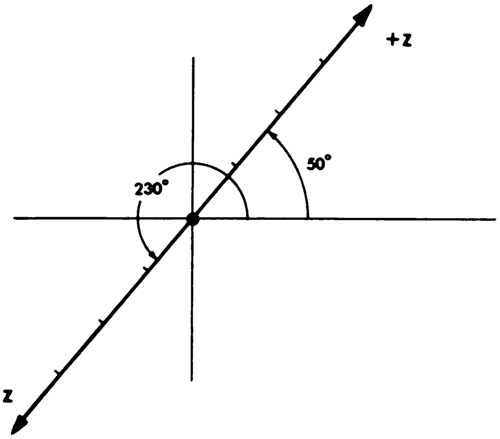
$$R=30\underline{/\ 0^\circ}$$

and $\quad -R=30\underline{/\ 180^\circ}$

$$Z=55\underline{/\ 58^\circ}$$



Adding Vectors on the DR Computer



But $X_L = Z + (-R)$

$$= 55\ /58^\circ + 30\ /180^\circ$$
$$= 46\ /90^\circ$$

Multiplication of Vectors

Occasionally, you may have to multiply and divide vector quantities. These operations cannot be done on the DR computer. However, the product of two vector quantities is found by *multiplying the magnitudes* of the vectors and *adding the angles* algebraically.

If you wish to multiply $Z_1 = 10/20^\circ$ and $Z_2 = 15/30^\circ$, proceed as follows:

The product of these two vectors is

$$(Z_1)\,(Z_2) = (10/20^\circ)\,(15/30^\circ)$$
$$= (10)\,(15)\ /20^\circ + 30^\circ$$
$$= 150/50^\circ$$

Division of Vectors

Division is the reverse process of multiplication. To divide two vectors then, *divide the magnitudes* of the two vectors and *subtract the angle* of the vector in the denominator from the angle of the vector in the numerator algebraically.

If you wish to divide Z_1 by Z_2 of the previous example, proceed as follows:

$$\frac{Z_1}{Z_2} = \frac{10/20^\circ}{15/30^\circ}$$
$$= \frac{10}{15}\ /20^\circ - 30^\circ$$
$$= .66\ / -10^\circ$$

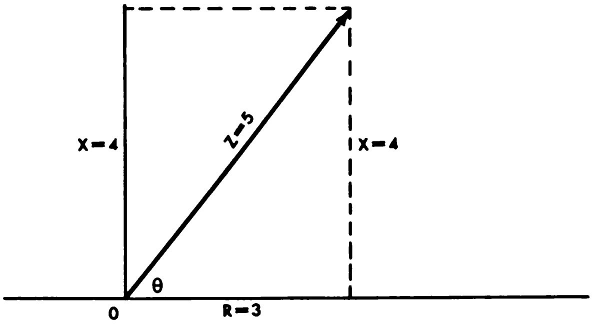
Converting from Polar to Rectangular Form

You have already seen on page 518, how complex quantities (such as *rectangular* forms of vectors) may be added and subtracted. Addition of vectors by graphical methods is not very accurate unless precision instruments are used. The accuracy of the DR computer varies, depending on the vectors to be added, and in some cases is only a close approximation. Whenever you wish to combine two vectors with absolute accuracy, convert the vectors to rectangular form and add them algebraically.

A vector may be represented in either *polar* or *rectangular* form. For example, 5/53° describes a vector 5 units in length at an angle of 53°. In the rectangular form, R+jx, the vector is resolved into two components. These components are the projections of the vector on the horizontal and vertical axes with their origin as the initial point of the vector.

The vector with its horizontal and vertical components are shown in the illustration below. The horizontal component is the *real* component and the vertical component is the *imaginary* or j component. To convert to rectangular form, you must find the vertical and horizontal components. Notice that the horizontal component R, is equal to Z cos θ, and that the vertical component X, is equal to Z sin θ. Thus 5/53° is equivalent to 5 cos θ+j 5 sin θ. Since cos 53° is .6 and sin 53° is .8,

$$5 \cos 53^\circ + j\,5 \sin 53^\circ = 5\,(.6) + j\,5\,(.8)$$
$$= 3 + j\,4$$



To convert 3+j 4 back from rectangular to polar form, you can see from the illustration that $\frac{X}{R} = \tan\theta$. Therefore, $\tan\theta = \frac{4}{3} = 1.3$ and

θ = arc tan 1.3 = 53°. Now that you know θ you can find Z as follows:

$$\cos\theta = \frac{R}{Z}$$

$$Z = \frac{R}{\cos\theta}$$

$$= \frac{3}{\cos 53°}$$

$$= \frac{3}{.6}$$

$$= 5$$

Therefore 3+j 4 represents $5\underline{/53°}$ in rectangular form.

To add vectors in polar form, convert to rectangular form and add algebraically. Convert the answer back to polar form. This procedure is shown in the following example.

Example

Add $35\underline{/40°}$ *and* $47\underline{/55°}$

Solution

$$35\underline{/40°} = 35 \cos 40° + j\,35 \sin 40$$

$$= 35\ (.7660) + j\,35\ (.6428)$$

$$= 26.81 + j\,22.50$$

$$47\underline{/55} = 47 \cos 55° + j\,47 \sin 55°$$

$$= 47\ (.5736) + j\,47\ (.8192)$$

$$= 26.96 + j\,38.5$$

Adding algebraically,

$$26.81 + j\,22.5$$

$$26.96 + j\,38.5$$

$$53.77 + j\,61$$

Converting this result to polar form,

$$\tan\theta = \frac{X}{R}$$

$$= \frac{61}{53.77}$$

$$= 1.134$$

$$\theta = \arctan 1.134 = 48.6°$$

$$Z = \frac{R}{\cos\theta}$$

$$= \frac{53.77}{\cos 48.6°}$$

$$= \frac{53.77}{.6613}$$

$$= 81.3$$

Therefore, the polar form of Z is,

$$Z = 81.3\underline{/48.6°}$$

Appendix

TRIGONOMETRIC TABLES

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
0.0	.00000	1.0000	.00000	Inf	90.0
.1	.00175	1.0000	.00175	573.0	.9
.2	.00349	1.0000	.00349	286.5	.8
.3	.00524	1.0000	.00524	191.0	.7
.4	.00698	1.0000	.00698	143.24	.6
.5	.00873	1.0000	.00873	114.59	.5
.6	.01047	0.9999	.01047	95.49	.4
.7	.01222	.9999	.01222	81.85	.3
.8	.01396	.9999	.01396	71.62	.2
.9	.01571	.9999	.01571	63.66	.1
1.0	.01745	0.9998	.01746	57.29	89.0
.1	.01920	.9998	.01920	52.80	.9
.2	.02094	.9998	.02095	47.74	.8
.3	.02269	.9997	.02269	44.07	.7
.4	.02443	.9997	.02444	40.92	.6
.5	.02618	.9997	.02619	38.19	.5
.6	.02792	.9996	.02793	35.80	.4
.7	.02967	.9996	.02968	33.69	.3
.8	.03141	.9995	.03143	31.82	.2
.9	.03316	.9995	.03317	30.14	.1
2.0	.03490	.9994	.03492	28.64	88.0
.1	.03664	.9993	.03667	27.27	.9
.2	.03839	.9993	.03842	26.03	.8
.3	.04013	.9992	.04016	24.90	.7
.4	.04188	.9991	.04191	23.86	.6
.5	.04362	.9990	.04366	22.90	.5
.6	.04536	.9990	.04541	22.02	.4
.7	.04711	.9989	.04716	21.20	.3
.8	.04885	.9988	.04891	20.45	.2
.9	.05059	.9987	.05066	19.74	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
3.0	.05234	0.9986	.05241	19.081	87.0
.1	.05408	.9985	.05416	18.464	.9
.2	.05582	.9984	.05591	17.886	.8
.3	.05756	.9983	.05766	17.343	.7
.4	.05931	.9982	.05941	16.832	.6
.5	.06105	.9981	.06116	16.350	.5
.6	.06279	.9980	.06291	15.895	.4
.7	.06453	.9979	.06467	15.464	.3
.8	.06627	.9978	.06642	15.056	.2
.9	.06802	.9977	.06817	14.669	.1
4.0	.06976	0.9976	.06993	14.301	86.0
.1	.07150	.9974	.07168	13.951	.9
.2	.07324	.9973	.07344	13.617	.8
.3	.07498	.9972	.07519	13.300	.7
.4	.07672	.9971	.07695	12.996	.6
.5	.07846	.9969	.07870	12.706	.5
.6	.08020	.9968	.08046	12.429	.4
.7	.08194	.9966	.08221	12.163	.3
.8	.08368	.9965	.08397	11.909	.2
.9	.08542	.9963	.08573	11.664	.1
5.0	.08716	0.9962	.08749	11.430	85.0
.1	.08889	.9960	.08925	11.205	.9
.2	.09063	.9959	.09101	10.988	.8
.3	.09237	.9957	.09277	10.780	.7
.4	.09411	.9956	.09453	10.579	.6
.5	.09585	.9954	.09629	10.385	.5
.6	.09758	.9952	.09805	10.199	.4
.7	.09932	.9951	.09981	10.019	.3
.8	.10106	.9949	.10158	9.845	.2
.9	.10279	.9947	.10334	9.677	.1
6.0	.10453	0.9945	.10510	9.514	84.0
.1	.10626	.9943	.10687	9.357	.9
.2	.10800	.9942	.10863	9.205	.8
.3	.10973	.9940	.11040	9.058	.7
.4	.11147	.9938	.11217	8.915	.6
.5	.11320	.9936	.11394	8.777	.5
.6	.11494	.9934	.11570	8.643	.4
.7	.11667	.9932	.11747	8.513	.3
.8	.11840	.9930	.11924	8.386	.2
.9	.12014	.9928	.12101	8.264	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
7.0	.12187	0.9925	.12278	8.144	83.0
.1	.12360	.9923	.12456	8.028	.9
.2	.12533	.9921	.12633	7.916	.8
.3	.12706	.9919	.12810	7.806	.7
.4	.12880	.9917	.12988	7.700	.6
.5	.13053	.9914	.13165	7.596	.5
.6	.13226	.9912	.13343	7.495	.4
.7	.13399	.9910	.13521	7.396	.3
.8	.13572	.9907	.13698	7.300	.2
.9	.13744	.9905	.13876	7.207	.1
8.0	.13917	0.9903	.14054	7.115	82.0
.1	.14090	.9900	.14232	7.026	.9
.2	14263	.9898	.14410	6.940	.8
.3	.14436	.9895	.14588	6.855	.7
.4	.14608	.9893	.14767	6.772	.6
.5	.14781	.9890	.14945	6.691	.5
.6	.14954	.9888	.15124	6.612	.4
.7	.15126	.9885	.15302	6.535	.3
.8	.15299	.9882	.15481	6.460	.2
.9	.15471	.9880	.15660	6.386	.1
9.0	.15643	0.9877	.15838	6.314	81.0
.1	.15816	.9874	.16017	6.243	.9
.2	.15988	.9871	.16196	6.174	.8
.3	.16160	.9869	.16376	6.107	.7
.4	.16333	.9866	.16555	6.041	.6
.5	.16505	.9863	.16734	5.976	.5
.6	.16677	.9860	.16914	5.912	.4
.7	.16849	.9857	.17093	5.850	.3
.8	.17021	.9854	.17273	5.789	.2
.9	.17193	.9851	.17453	5.730	.1
10.0	.1736	0.9848	.1763	5.671	80.0
.1	.1754	.9845	.1781	5.614	.9
.2	.1771	.9842	.1799	5.558	.8
.3	.1788	.9839	.1817	5.503	.7
.4	.1805	.9836	.1835	5.499	.6
.5	.1822	.9833	.1853	5.396	.5
.6	.1840	.9829	.1871	5.343	.4
.7	.1857	.9826	.1890	5.292	.3
.8	.1874	.9823	.1908	5.242	.2
.9	.1891	.9820	.1926	5.193	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
11.0	.1908	0.9816	.1944	5.145	79.0
.1	.1925	.9813	.1962	5.097	.9
.2	.1942	.9810	.1980	5.050	.8
.3	.1959	.9806	.1998	5.005	.7
.4	.1977	.9803	.2016	4.959	.6
.5	.1994	.9799	.2035	4.915	.5
.6	.2011	.9796	.2053	4.872	.4
.7	.2028	.9792	.2071	4.829	.3
.8	.2045	.9789	.2089	4.787	.2
.9	.2062	.9785	.2107	4.745	.1
12.0	0.2079	0.9781	0.2126	4.705	78.0
.1	.2096	.9778	.2144	4.665	.9
.2	.2113	.9774	.2162	4.625	.8
.3	.2130	.9770	.2180	4.586	.7
.4	.2147	.9767	.2199	4.548	.6
.5	.2164	.9763	.2217	4.511	.5
.6	.2181	.9759	.2235	4.474	.4
.7	.2198	.9755	.2254	4.437	.3
.8	.2215	.9751	.2272	4.402	.2
.9	.2233	.9748	.2290	4.366	.1
13.0	0.2250	0.9744	0.2309	4.331	77.0
.1	.2267	.9740	.2327	4.297	.9
.2	.2284	.9736	.2345	4.264	.8
.3	.2300	.9732	.2364	4.230	.7
.4	.2317	.9728	.2382	4.198	.6
.5	.2334	.9724	.2401	4.165	.5
.6	.2351	.9720	.2419	4.134	.4
.7	.2368	.9715	.2438	4.102	.3
.8	.2385	.9711	.2456	4.071	.2
.9	.2402	.9707	.2475	4.041	.1
14.0	0.2419	0.9703	0.2493	4.011	76.0
.1	.2436	.9699	.2512	3.981	.9
.2	.2453	.9694	.2530	3.952	.8
.3	.2470	.9690	.2549	3.923	.7
.4	.2487	.9686	.2568	3.895	.6
.5	.2504	.9681	.2586	3.867	.5
.6	.2521	.9677	.2605	3.839	.4
.7	.2538	.9673	.2623	3.812	.3
.8	.2554	.9668	.2642	3.785	.2
.9	.2571	.9664	.2661	3.758	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
15.0	0.2588	0.9659	0.2679	3.732	75.0
.1	.2605	.9655	.2698	3.706	.9
.2	.2622	.9650	.2717	3.681	.8
.3	.2639	.9646	.2736	3.655	.7
.4	.2656	.9641	.2754	3.630	.6
.5	.2672	.9636	.2773	3.606	.5
.6	.2689	.9632	.2792	3.582	.4
.7	.2706	.9627	.2811	3.558	.3
.8	.2723	.9622	.2830	3.534	.2
.9	.2740	.9617	.2849	3.511	.1
16.0	0.2756	0.9613	0.2867	3.487	74.0
.1	.2773	.9608	.2886	3.465	.9
.2	.2790	.9603	.2905	3.442	.8
.3	.2807	.9598	.2924	3.420	.7
.4	.2823	.9593	.2943	3.398	.6
.5	.2840	.9588	.2962	3.376	.5
.6	.2857	.9583	.2981	3.354	.4
.7	.2874	.9578	.3000	3.333	.3
.8	.2890	.9573	.3019	3.312	.2
.9	.2907	.9568	.3038	3.291	.1
17.0	0.2924	0.9563	0.3057	3.271	73.0
.1	.2940	.9558	.3076	3.251	.9
.2	.2957	.9553	.3096	3.230	.8
.3	.2974	.9548	.3115	3.211	.7
.4	.2990	.9542	.3134	3.191	.6
.5	.3007	.9537	.3153	3.172	.5
.6	.3024	.9532	.3172	3.152	.4
.7	.3040	.9527	.3191	3.133	.3
.8	.3057	.9521	.3211	3.115	.2
.9	.3074	.9516	.3230	3.096	.1
18.0	0.3090	0.9511	0.3249	3.078	72.0
.1	.3107	.9505	.3269	3.060	.9
.2	.3123	.9500	.3288	3.042	.8
.3	.3140	.9494	.3307	3.024	.7
.4	.3156	.9489	.3327	3.006	.6
.5	.3173	.9483	.3346	2.989	.5
.6	.3190	.9478	.3365	2.971	.4
.7	.3206	.9472	.3385	2.954	.3
.8	.3223	.9466	.3404	2.937	.2
.9	.3239	.9461	.3424	2.921	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
19.0	0.3256	0.9455	0.3443	2.904	71.0
.1	.3272	.9449	.3463	2.888	.9
.2	.3289	.9444	.3482	2.872	.8
.3	.3305	.9438	.3502	2.856	.7
.4	.3322	.9432	.3522	2.840	.6
.5	.3338	.9426	.3541	2.824	.5
.6	.3355	.9421	.3561	2.808	.4
.7	.3371	.9415	.3581	2.793	.3
.8	.3387	.9409	.3600	2.778	.2
.9	.3404	.9403	.3620	2.762	.1
20.0	0.3420	0.9397	0.3640	2.747	70.0
.1	.3437	.9391	.3659	2.733	.9
.2	.3453	.9385	.3679	2.718	.8
.3	.3469	.9379	.3699	2.703	.7
.4	.3486	.9373	.3719	2.689	.6
.5	.3502	.9367	.3739	2.675	.5
.6	.3518	.9361	.3759	2.660	.4
.7	.3535	.9354	.3779	2.646	.3
.8	.3551	.9348	.3799	2.633	.2
.9	.3567	.9342	.3819	2.619	.1
21.0	0.3584	0.9336	0.3839	2.605	69.0
.1	.3600	.9330	.3859	2.592	.9
.2	.3616	.9323	.3879	2.578	.8
.3	.3633	.9317	.3899	2.565	.7
.4	.3649	.9311	.3919	2.552	.6
.5	.3665	.9304	.3939	2.539	.5
.6	.3681	.9298	.3959	2.526	.4
.7	.3697	.9291	.3979	2.513	.3
.8	.3714	.9285	.4000	2.500	.2
.9	.3730	.9278	.4020	2.488	.1
22.0	0.3746	0.9272	0.4040	2.475	68.0
.1	.3762	.9265	.4061	2.463	.9
.2	.3778	.9259	.4081	2.450	.8
.3	.3795	.9252	.4101	2.438	.7
.4	.3811	.9245	.4122	2.426	.6
.5	.3827	.9239	.4142	2.414	.5
.6	.3843	.9232	.4163	2.402	.4
.7	.3859	.9225	.4183	2.391	.3
.8	.3875	.9219	.4204	2.379	.2
.9	.3891	.9212	.4224	2.367	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
23.0	0.3907	0.9205	0.4245	2.356	67.0
.1	.3923	.9198	.4265	2.344	.9
.2	.3939	.9191	.4286	2.333	.8
.3	.3955	.9184	.4307	2.322	.7
.4	.3971	.9178	.4327	2.311	.6
.5	.3987	.9171	.4348	2.300	.5
.6	.4003	.9164	.4369	2.289	.4
.7	.4019	.9157	.4390	2.278	.3
.8	.4035	.9150	.4411	2.267	.2
.9	.4051	.9143	.4431	2.257	.1
24.0	0.4067	0.9135	0.4452	2.246	66.0
.1	.4083	.9128	.4473	2.236	.9
.2	.4099	.9121	.4494	2.225	.8
.3	.4115	.9114	.4515	2.215	.7
.4	.4131	.9107	.4536	2.204	.6
.5	.4147	.9100	.4557	2.194	.5
.6	.4163	.9092	.4578	2.184	.4
.7	.4179	.9085	.4599	2.174	.3
.8	.4195	.9078	.4621	2.164	.2
.9	.4210	.9070	.4642	2.154	.1
25.0	0.4226	0.9063	0.4663	2.145	65.0
.1	.4242	.9056	.4684	2.135	.9
.2	.4258	.9048	.4706	2.125	.8
.3	.4274	.9041	.4727	2.116	.7
.4	.4289	.9033	.4748	2.106	.6
.5	.4305	.9026	.4770	2.097	.5
.6	.4321	.9018	.4791	2.087	.4
.7	.4337	.9011	.4813	2.078	.3
.8	.4352	.9003	.4834	2.069	.2
.9	.4368	.8996	.4856	2.059	.1
26.0	0.4384	0.8988	0.4877	2.050	64.0
.1	.4399	.8980	.4899	2.041	.9
.2	.4415	.8973	.4921	2.032	.8
.3	.4431	.8965	.4942	2.023	.7
.4	.4446	.8957	.4964	2.014	.6
.5	.4462	8949	.4986	2.006	.5
.6	.4478	.8942	.5008	1.997	.4
.7	.4493	.8934	.5029	1.988	.3
.8	.4509	.8926	.5051	1.980	.2
.9	.4524	.8918	.5073	1.971	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
27.0	0.4540	0.8910	0.5095	1.963	63.0
.1	.4555	.8902	.5117	1.954	.9
.2	.4571	.8894	.5139	1.946	.8
.3	.4586	.8886	.5161	1.937	.7
.4	.4602	.8878	.5184	1.929	.6
.5	.4617	.8870	.5206	1.921	.5
.6	.4633	.8862	.5228	1.913	.4
.7	.4648	.8854	.5250	1.905	.3
.8	.4664	.8846	.5272	1.897	.2
.9	.4679	.8838	.5295	1.889	.1
28.0	0.4695	0.8829	0.5317	1.881	62.0
.1	.4710	.8821	.5340	1.873	.9
.2	.4726	.8813	.5362	1.865	.8
.3	.4741	.8805	.5384	1.857	.7
.4	.4756	.8796	.5407	1.849	.6
.5	.4772	.8788	.5430	1.842	.5
.6	.4787	.8780	.5452	1.834	.4
.7	.4802	.8771	.5475	1.827	.3
.8	.4818	.8763	.5498	1.819	.2
.9	.4833	.8755	.5520	1.811	.1
29.0	0.4848	0.8746	0.5543	1.804	61.0
.1	.4863	.8738	.5566	1.797	.9
.2	.4879	.8729	.5589	1.789	.8
.3	.4894	.8721	.5612	1.782	.7
.4	.4909	.8712	.5635	1.775	.6
.5	.4924	.8704	.5658	1.767	.5
.6	.4939	.8695	.5681	1.760	.4
.7	.4955	.8686	.5704	1.753	.3
.8	.4970	.8678	.5727	1.746	.2
.9	.4985	.8669	.5750	1.739	.1
30.0	0.5000	0.8660	0.5774	1.7321	60.0
.1	.5015	.8652	.5797	1.7251	.9
.2	.5030	.8643	.5820	1.7182	.8
.3	.5045	.8634	.5844	1.7113	.7
.4	.5060	.8625	.5867	1.7045	.6
.5	.5075	.8616	.5890	1.6977	.5
.6	.5090	.8607	.5914	1.6909	.4
.7	.5105	.8599	.5938	1.6842	.3
.8	.5120	.8590	.5961	1.6775	.2
.9	.5135	.8581	.5985	1.6709	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG	DEGREE
31.0	0.5150	0.8572	0.6009	1.6643	59.0
.1	.5165	.8563	.6032	1.6577	.9
.2	.5180	.8554	.6056	1.6512	.8
.3	.5195	.8545	.6080	1.6447	.7
.4	.5210	.8536	.6104	1.6383	.6
.5	.5225	.8526	.6128	1.6319	.5
.6	.5240	.8517	.6152	1.6255	.4
.7	.5255	.8508	.6176	1.6191	.3
.8	.5270	.8499	.6200	1.6128	.2
.9	.5284	.8490	.6224	1.6066	.1
32.0	0.5299	0.8480	0.6249	1.6003	58.0
.1	.5314	.8471	.6273	1.5941	.9
.2	.5329	.8462	.6297	1.5880	.8
.3	.5344	.8453	.6322	1.5818	.7
.4	.5358	.8443	.6346	1.5757	.6
.5	.5373	.8434	.6371	1.5697	.5
.6	.5388	.8425	.6395	1.5637	.4
.7	.5402	.8415	.6420	1.5577	.3
.8	.5417	.8406	.6445	1.5517	.2
.9	.5432	.8396	.6469	1.5458	.1
33.0	0.5446	0.8387	0.6494	1.5399	57.0
.1	.5461	.8377	.6519	1.5340	.9
.2	.5476	.8368	.6544	1.5282	.8
.3	.5490	.8358	.6569	1.5224	.7
.4	.5505	.8348	.6594	1.5166	.6
.5	.5519	.8339	.6619	1.5108	.5
.6	.5534	.8329	.6644	1.5051	.4
.7	.5548	.8320	.6669	1.4994	.3
.8	.5563	.8310	.6694	1.4938	.2
.9	.5577	.8300	.6720	1.4882	.1
34.0	0.5592	0.8290	0.6745	1.4826	56.0
.1	.5606	.8281	.6771	1.4770	.9
.2	.5621	.8271	.6796	1.4715	.8
.3	.5635	.8261	.6822	1.4659	.7
.4	.5650	.8251	.6847	1.4605	.6
.5	.5664	.8241	.6873	1.4550	.5
.6	.5678	.8231	.6899	1.4496	.4
.7	.5693	.8221	.6924	1.4442	.3
.8	.5707	.8211	.6950	1.4388	.2
.9	.5721	.8202	.6976	1.4335	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
35.0	0.5736	0.8192	0.7002	1.4281	55.0
.1	.5750	.8181	.7028	1.4229	.9
.2	.5764	.8171	.7054	1.4176	.8
.3	.5779	.8161	.7080	1.4124	.7
.4	.5793	.8151	.7107	1.4071	.6
.5	.5807	.8141	.7133	1.4019	.5
.6	.5821	.8131	.7159	1.3968	.4
.7	.5835	.8121	.7186	1.3916	.3
.8	.5850	.8111	.7212	1.3865	.2
.9	.5864	.8100	.7239	1.3814	.1
36.0	0.5878	0.8090	0.7265	1.3764	54.0
.1	.5892	.8080	.7292	1.3713	.9
.2	.5906	.8070	.7319	1.3663	.8
.3	.5920	.8059	.7346	1.3613	.7
.4	.5934	.8049	.7373	1.3564	.6
.5	.5948	.8039	.7400	1.3514	.5
.6	.5962	.8028	.7427	1.3465	.4
.7	.5976	.8018	.7454	1.3416	.3
.8	.5990	.8007	.7481	1.3367	.2
.9	.6004	.7997	.7508	1.3319	.1
37.0	0.6018	0.7986	0.7536	1.3270	53.0
.1	.6032	.7976	.7563	1.3222	.9
.2	.6046	.7965	.7590	1.3175	.8
.3	.6060	.7955	.7618	1.3127	.7
.4	.6074	.7944	.7646	1.3079	.6
.5	.6088	.7934	.7673	1.3032	.5
.6	.6101	.7923	.7701	1.2985	.4
.7	.6115	.7912	.7729	1.2938	.3
.8	.6129	.7902	.7757	1.2892	.2
.9	.6143	.7891	.7785	1.2846	.1
38.0	0.6157	0.7880	0.7813	1.2799	52.0
.1	.6170	.7869	.7841	1.2753	.9
.2	.6184	.7859	.7869	1.2708	.8
.3	.6198	.7848	.7898	1.2662	.7
.4	.6211	.7837	.7926	1.2617	.6
.5	.6225	.7826	.7954	1.2572	.5
.6	.6239	.7815	.7983	1.2527	.4
.7	.6252	.7804	.8012	1.2482	.3
.8	.6266	.7793	.8040	1.2437	.2
.9	.6280	.7782	.8069	1.2393	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
39.0	0.6293	0.7771	0.8098	1.2349	51.0
.1	.6307	.7760	.8127	1.2305	.9
.2	.6320	.7749	.8156	1.2261	.8
.3	.6334	.7738	.8185	1.2218	.7
.4	.6347	.7727	.8214	1.2174	.6
.5	.6361	.7716	.8243	1.2131	.5
.6	.6374	.7705	.8273	1.2088	.4
.7	.6388	.7694	.8302	1.2045	.3
.8	.6401	.7683	.8332	1.2002	.2
.9	.6414	.7672	.8361	1.1960	.1
40.0	0.6428	0.7660	0.8391	1.1918	50.0
.1	.6441	.7649	.8421	1.1875	.9
.2	.6455	.7638	.8451	1.1833	.8
.3	.6468	.7627	.8481	1.1792	.7
.4	.6481	.7615	.8511	1.1750	.6
.5	.6494	.7604	.8541	1.1708	.5
.6	.6508	.7593	.8571	1.1667	.4
.7	.6521	.7581	.8601	1.1626	.3
.8	.6534	.7570	.8632	1.1585	.2
.9	.6547	.7559	.8662	1.1544	.1
41.0	0.6561	0.7547	0.8693	1.1504	49.0
.1	.6574	.7536	.8724	1.1463	.9
.2	.6587	.7524	.8754	1.1423	.8
.3	.6600	.7513	.8785	1.1383	.7
.4	.6613	.7501	.8816	1.1343	.6
.5	.6626	.7490	.8847	1.1303	.5
.6	.6639	.7478	.8878	1.1263	.4
.7	.6652	.7466	.8910	1.1224	.3
.8	.6665	.7455	.8941	1.1184	.2
.9	.6678	.7443	.8972	1.1145	.1
42.0	0.6691	0.7431	0.9004	1.1106	48.0
.1	.6704	.7420	.9036	1.1067	.9
.2	.6717	.7408	.9067	1.1028	.8
.3	.6730	.7396	.9099	1.0990	.7
.4	.6743	.7385	.9131	1.0951	.6
.5	.6756	.7373	.9163	1.0913	.5
.6	.6769	.7361	.9195	1.0875	.4
.7	.6782	.7349	.9228	1.0837	.3
.8	.6794	.7337	.9260	1.0799	.2
.9	.6807	.7325	.9293	1.0761	.1
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

DEGREE	SINE	COSINE	TANGENT	COTANG.	DEGREE
43.0	0.6820	0.7314	0.9325	1.0724	47.0
.1	.6833	.7302	.9358	1.0686	.9
.2	.6845	.7290	.9391	1.0649	.8
.3	.6858	.7278	.9424	1.0612	.7
.4	.6871	.7266	.9457	1.0575	.6
.5	.6884	.7254	.9490	1.0538	.5
.6	.6896	.7242	.9523	1.0501	.4
.7	.6909	.7230	.9556	1.0464	.3
.8	.6921	.7218	.9590	1.0428	.2
.9	.6934	.7206	.9623	1.0392	.1
44.0	0.6947	0.7193	0.9657	1.0355	46.0
.1	.6959	.7181	.9691	1.0319	.9
.2	.6972	.7169	.9725	1.0283	.8
.3	.6984	.7157	.9759	1.0247	.7
.4	.6997	.7145	.9793	1.0212	.6
.5	.7009	.7133	.9827	1.0176	.5
.6	.7022	.7120	.9861	1.0141	.4
.7	.7034	.7108	.9896	1.0105	.3
.8	.7046	.7096	.9930	1.0070	.2
.9	.7059	.7083	.9965	1.0035	.1
45.0	0.7071	0.7071	1.0000	1.0000	45.0
DEGREE	COSINE	SINE	COTANG.	TANGENT	DEGREE

FOUR PLACE COMMON LOGARITHM TABLES

N.	0	1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404
22	3424	3444	3464	3483	3502	3522	3541	3460	3579	3598
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3984
25	3979	3997	4010	4031	4048	4065	4082	4099	4116	4133
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117
41	6128	6138	6149	6160	6170	5180	6191	6201	6212	6222
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396
N.	0	1	2	3	4	5	6	7	8	9

FOUR PLACE COMMON LOGARITHM TABLES

N.	0	1	2	3	4	5	6	7	8	9
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055
64	8062	8069	8075	8082	8089	8096	8102	8109	8118	8122
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319
68	8325	8331	8338	8344	8351	8357	8363	8370	8386	8382
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567
72	8573	8579	8585	8591	8597	8603	8603	8615	8621	8627
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633
92	9638	9643	9647	9653	9657	9661	9666	9671	9675	9680
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996
N.	0	1	2	3	4	5	6	7	8	9

SQUARES, CUBES AND ROOTS

Roots of numbers other than those given directly may be found by the following relations:

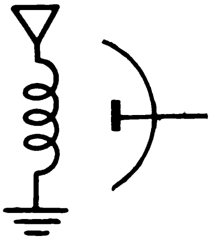
$\sqrt{100n}=10\sqrt{n}$; $\sqrt{1000n}=10\sqrt{10n}$; $\sqrt{\frac{1}{10}n}=\frac{1}{10}\sqrt{10n}$; $\sqrt{\frac{1}{100}n}=\frac{1}{10}\sqrt{n}$; $\sqrt{\frac{1}{1000}n}=\frac{1}{100}\sqrt{10n}$; $\sqrt[3]{1000n}=10\sqrt[3]{n}$; $\sqrt[3]{10{,}000n}=10\sqrt[3]{10n}$; $\sqrt[3]{100{,}000n}=10\sqrt[3]{100n}$; $\sqrt[3]{\frac{1}{10}n}=\frac{1}{10}\sqrt[3]{100n}$; $\sqrt[3]{\frac{1}{100}n}=\frac{1}{10}\sqrt[3]{10n}$; $\sqrt[3]{\frac{1}{1000}n}=\frac{1}{10}\sqrt[3]{n}$.

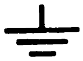
n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n^3	$\sqrt[3]{n}$	$\sqrt[3]{10n}$	$\sqrt[3]{100n}$
1	1	1.000 000	3.162 278	1	1.000 000	2.154 435	4.641 589
2	4	1.414 214	4.472 136	8	1.259 921	2.714 418	5.848 035
3	9	1.732 051	5.477 226	27	1.442 250	3.107 233	6.694 330
4	16	2.000 000	6.324 555	64	1.587 401	3.419 952	7.368 063
5	25	2.236 068	7.071 068	125	1.709 976	3.634 031	7.937 005
6	36	2.449 490	7.745 967	216	1.817 121	3.914 868	8.434 327
7	49	2.645 751	8.366 600	343	1.912 931	4.121 285	8.879 040
8	64	2.828 427	8.944 272	512	2.000 000	4.308 869	9.283 178
9	81	3.000 000	9.486 833	729	2.080 084	4.481 405	9.654 894
10	100	3.162 278	10.00000	1 000	2.154 435	4.641 589	10.00000
11	121	3.316 625	10.48800	1 331	2.223 980	4.791 420	10.32280
12	144	3.464 102	10.95145	1 728	2.280 428	4.932 424	10.62659
13	169	3.605 551	11.40175	2 197	2.351 335	5.065 797	10.91393
14	196	3.741 657	11.83216	2 744	2.410 142	5.192 494	11.18689
15	225	3.872 983	12.24745	3 375	2.466 212	5.313 293	11.44714
16	256	4.000 000	12.64911	4 096	2.519 842	5.428 835	11.69607
17	289	4.123 106	13.03840	4 913	2.571 282	5.539 658	11.93483
18	324	4.242 641	13.41641	5 832	2.620 741	5.646 216	12.16440
19	361	4.358 899	13.78405	6 859	2.668 402	5.748 897	12.38562
20	400	4.472 136	14.14214	8 000	2.714 418	5.848 035	12.59921
21	441	4.582 576	14.49138	9 261	2.758 924	5.943 922	12.80579
22	484	4.690 416	14.83240	10 648	2.802 039	6.036 811	13.00591
23	529	4.705 832	15.16575	12 167	2.843 867	6.126 926	13.20006
24	576	4.898 979	15.49193	13 824	2.884 499	6.214 465	13.38866
25	625	5.000 000	15.81139	15 625	2.924 018	6.299 605	13.57209
26	676	5.099 020	16.12452	17 576	2.962 496	6.382 504	13.75069
27	729	5.196 152	16.43168	19 683	3.000 000	6.463 304	13.92477
28	784	5.291 503	16.73320	21 952	3.036 589	6.542 133	14.09460
29	841	5.385 165	17.02939	24 389	3.072 317	6.619 106	14.26043
30	900	5.477 226	17.32051	27 000	3.107 233	6.694 330	14.42250
31	961	5.567 674	17.60682	29 791	3.141 381	6.767 899	14.58100
32	1 024	5.656 854	17.88854	32 768	3.174 802	6.839 904	14.73613
33	1 089	5.744 563	18.16590	35 937	3.207 534	6.910 423	14.88806
34	1 156	5.830 952	18.43909	39 304	3.239 612	6.979 532	15.03695
35	1 225	5.916.080	18.70829	42 875	3.271 066	7.047 299	15.18294
36	1 296	6.000 000	18.97367	46 656	3.301 927	7.113 787	15.32619
37	1 389	6.082 763	19.23538	50 653	3.332 222	7.179 054	15.46680
38	1 444	6.164 414	19.49359	54 872	3.361 975	7.243 156	15.60491
39	1 521	6.244 998	19.74842	59 319	3.391 211	7.306 144	15.74061
40	1 600	6.324 555	20.00000	64 000	3.419 952	7.368 063	15.87401
41	1 681	6.403 124	20.24846	68 921	3.448 217	7.428 959	16.00521
42	1 764	6.430 741	20.49390	74 088	3.476 027	7.488 872	16.13429
43	1 849	6.557 439	20.73644	79 507	3.503 398	7.547 842	16.26133
44	1 936	6.633 250	20.97618	85 184	3.530 348	7.605 905	16.38643
45	2 025	6.708 204	21.21320	91 125	3.556 893	7.663 094	16.50964
46	2 116	6.782 330	21.44761	97 336	3.583 048	7.719 443	16.63103
47	2 209	6.855 655	21.67948	103 823	3.608 826	7.774 980	16.75069
48	2 304	6.928 203	21.90890	110 592	3.634 241	7.829 735	16.86865
49	2 401	7.000 000	22.13594	117 649	3.659 306	7.883 735	16.98499
50	2 500	7.071 068	22.36068	125 000	3.684 031	7.937 005	17.00076

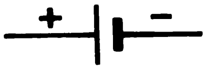
SQUARES, CUBES AND ROOTS (Continued)

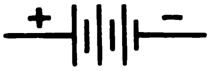
n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n^3	$\sqrt[3]{n}$	$\sqrt[3]{10n}$	$\sqrt[3]{100b}$
50	2 500	7.071 068	22.36068	125 000	3.684 031	7.937 005	17.09976
51	2 601	7.141 428	22.58318	132 651	3.708 430	7.989 570	17.21301
52	2 704	7.211 103	22.80351	140 608	3.732 511	8.041 452	17.32478
53	2 809	7.280 110	23.02173	148 877	3.756 286	8.092 672	17.43513
54	2 916	7.348 469	23.23790	157 464	3.779 763	8.143 253	17.54411
55	3 025	7.416 198	23.45208	166 375	3.802 952	8.193 213	17.65174
56	3 136	7.483 315	23.66432	175 616	3.825 862	8.242 571	17.75808
57	3 249	7.549 834	23.87467	185 193	3.848 501	8.291 344	17.86316
58	3 364	7.615 773	24.08319	195 112	3.870 877	8.339 551	17.96702
59	3 481	7.681 146	24.28992	205 379	3.892 996	8.387 207	18.06969
60	3 600	7.745 967	24.49490	216 000	3.914 868	8.434 327	18.17121
61	3 721	7.810 250	24.69818	226 981	3.936 497	8.480 926	18.27160
62	3 844	7.874 008	24.89980	238 328	3.957 892	8.527 019	18.37091
63	3 969	7.937 254	25.09980	250 047	3.979 057	8.572 619	18.46915
64	4 096	8.000 000	25.29822	262 144	4.000 000	8.617 739	18.56636
65	4 225	8.062 258	25.49510	274 625	4.020 726	8.662 391	18.66256
66	4 356	8.124 038	25.69047	287 496	4.041 240	8.706 588	18.75777
67	4 489	8.185 353	25.88436	300 763	4.061 548	8.750 340	18.85204
68	4 624	8.246 211	26.07681	314 432	4.081 655	8.793 659	18.94536
69	4 761	8.306 624	26.26785	328 509	4.101 566	8.836 556	19.03778
70	4 900	8.366 600	26.45751	343 000	4.121 285	8.879 040	19.12931
71	5 041	8.426 150	26.64583	357 911	4.140 818	8.921 121	19.21997
72	5 184	8.485 281	26.83282	373 248	4.160 168	8.962 809	19.30979
73	5 329	8.544 004	27.01851	389 017	4.179 339	9.004 113	19.39877
74	5 476	8.602 325	27.20294	405 224	4.198 336	9.045 042	19.48695
75	5 625	8.660 254	27.38613	421 875	4.217 163	9.058 603	19.57434
76	5 776	8.717 798	27.56810	438 976	4.235 824	9.125 805	19.66095
77	5 929	8.774 964	27.74887	456 533	4.254 321	9.165 656	19.74681
78	6 084	8.831 761	27.92848	474 552	4.272 659	9.205 164	19.83192
79	6 241	8.888 194	28.10694	493 039	4.290 840	9.244 335	19.91632
80	6 400	8.944 272	28.28427	512 000	4.308 869	9.283 178	20.00000
81	6 561	9.000 000	28.46050	531 441	4.326 749	9.321 698	20.08299
82	6 724	9.055 385	28.63564	551 368	4.344 481	9.359 902	20.16530
83	6 889	9.110 434	28.80972	571 787	4.362 071	9.397 796	20.24694
84	7 056	9.165 151	28.98275	592 704	4.379 519	9.435 388	20.32793
85	7 225	9.219 544	29.15476	614 125	4.396 830	9.472 682	20.40828
86	7 396	9.273 618	29.32576	636 056	4.414 005	9.509 685	20.48800
87	7 569	9.327 379	29.49576	658 503	4.431 048	9.546 403	20.56710
88	7 744	9.380 832	29.66479	681 472	4.447 960	9.582 840	20.61560
89	7 921	9.433 981	29.83287	704 969	4.464 745	9.619 002	20.72351
90	8 100	9.486 833	30.00000	729 000	4.481 405	9.654 894	20.80084
91	8 281	9.539 392	30.16621	753 571	4.497 941	9.690 521	20.87759
92	8 464	9.591 663	30.33150	778 688	4.514 357	9.725 888	20.95379
93	8 649	9.643 651	30.49590	804 357	4.530 655	9.761 000	21.02944
94	8 836	9.695 360	30.65942	830 584	4.546 836	9.795 861	21.10454
95	9 025	9.746 794	30.82207	857 375	4.562 903	9.830 476	21.17912
96	9 216	9.797 959	30.98387	884 736	4.578 857	9.864 848	21.25317
97	9 409	9.848 858	31.14482	912 673	4.594 701	9.898 983	21.32671
98	9 604	9.899 495	31.30495	941 192	4.610 436	9.932 884	21.39975
99	9 801	9.949 874	31.46427	970 299	4.626 065	9.966 555	21.47229
100	10 000	10.00000	31.62278	1 000 000	4.641 589	10.00000	21.54435

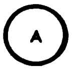
symbols for diagrams


ANTENNAS

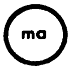

GROUND


BATTERY CELL

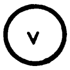

BATTERY OF THREE CELLS



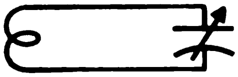
AMMETER

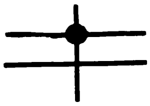


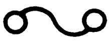
MILLIAMMETER

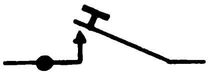


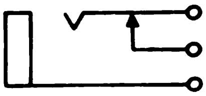
VOLTMETER

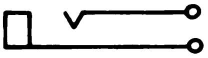

FREQUENCY METER OR WAVE METER

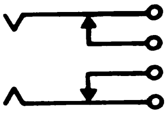

CONDUCTORS JOINED
CONDUCTORS NOT JOINED

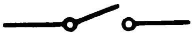

FUSE

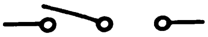

KEY

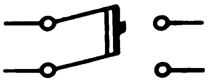

CLOSED CIRCUIT JACK

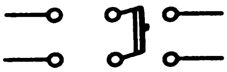

OPEN CIRCUIT JACK

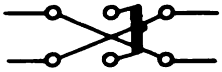

MULTIPLE CIRCUIT JACK

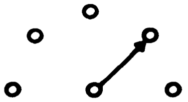

SINGLE-POLE, SINGLE-THROW SWITCH

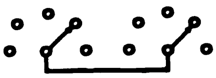

SINGLE-POLE, DOUBLE-THROW SWITCH

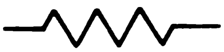

DOUBLE-POLE, SINGLE-THROW SWITCH

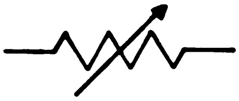

DOUBLE-POLE, DOUBLE-THROW SWITCH

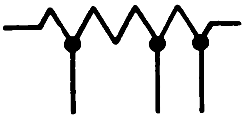

DOUBLE-POLE, DOUBLE-THROW, REVERSING SW.


SINGLE-DECK, CIRCUIT-SELECTOR SWITCH

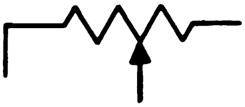

MULTIPLE-DECK, CIRCUIT-SELECTOR SWITCH


FIXED RESISTOR

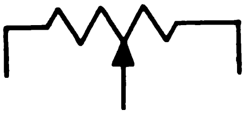

ADJUSTABLE RESISTOR



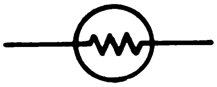
TAPPED RESISTOR



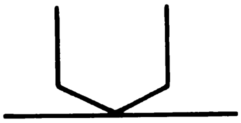
RHEOSTAT



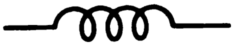
POTENTIOMETER



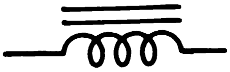
BALLAST
RESISTOR



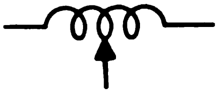
THERMO
ELEMENT



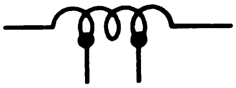
INDUCTOR,
AIR-CORE



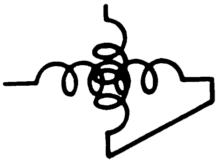
INDUCTOR,
IRON-CORE



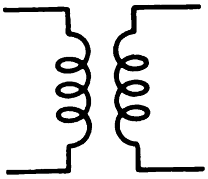
ADJUSTABLE
INDUCTOR



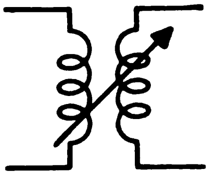
TAPPED
INDUCTOR



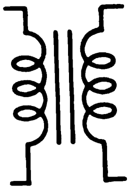
VARIABLE INDUCTOR
(VARIOMETER)



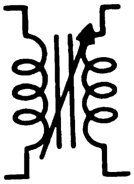
TRANSFORMER,
AIR-CORE



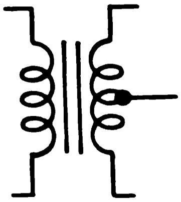
TRANSFORMER WITH
VARIABLE COUPLING
(VARIOCOUPLER WITH
MOVING COIL
INDICATOR)



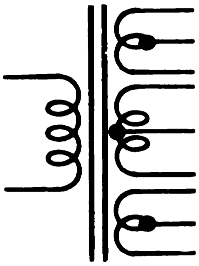
TRANSFORMER,
IRON-CORE



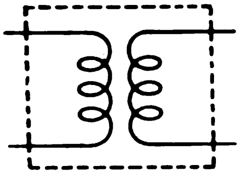
VARIABLE-CORE
TRANSFORMER



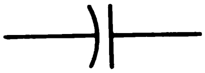
PUSH-PULL
TRANSFORMER



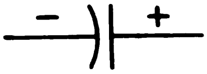
POWER
TRANSFORMER



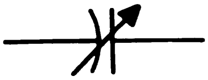
SHIELDED TRANSFORMER



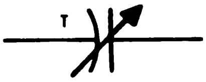
FIXED CAPACITOR (PAPER, MICA OR OIL)



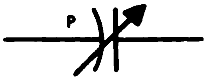
ELECTROLYTIC CAPACITOR



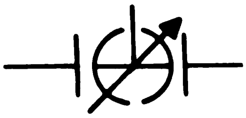
VARIABLE CAPACITOR



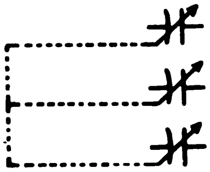
ADJUSTABLE CAPACITOR (TRIMMER)



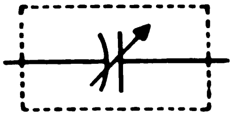
ADJUSTABLE CAPACITOR (PADDER)



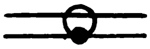
SPLIT STATOR VARIABLE CAPACITOR



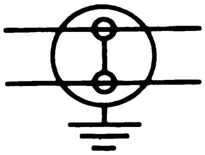
GANGED VARIABLE CAPACITORS MECHANICAL LINKAGE



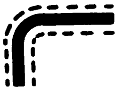
SHIELDED CAPACITOR



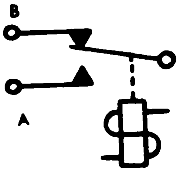
COAXIAL CABLE



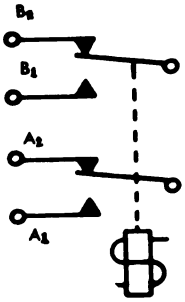
TWIN COAXIAL CABLE



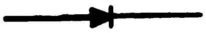
SHIELDED CABLE



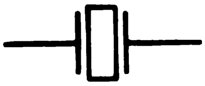
RELAY; CIRCUIT A OPEN WHEN DE-ENERGIZED



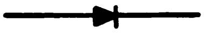
RELAY; CIRCUITS A_2 AND B_2 CLOSED WHEN DE-ENERGIZED



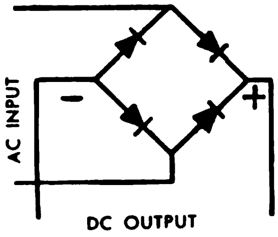
CRYSTAL DETECTOR



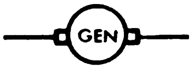
CRYSTAL



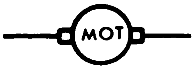
CONTACT RECTIFIER, HALF-WAVE



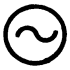
CONTACT RECTIFIER, FULL-WAVE



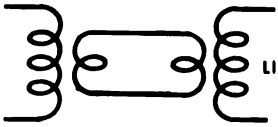
GENERATOR



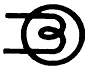
MOTOR



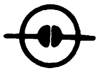
A C VOLTAGE SOURCE



LINK COUPLING



PILOT LAMP



NEON LAMP

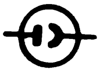
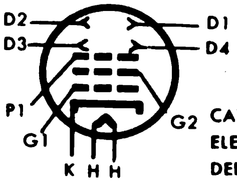
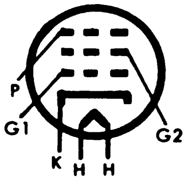


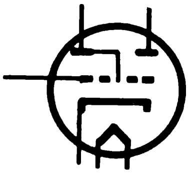
PHOTO TUBE



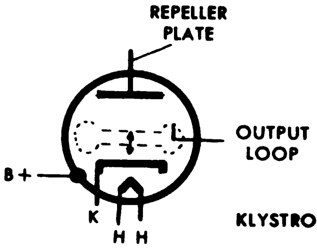
CATHODE-RAY TUBE, ELECTROSTATIC-DEFLECTION



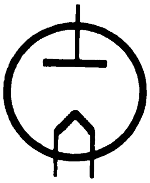
CATHODE-RAY TUBE, ELECTRO-MAGNETIC-DEFLECTION



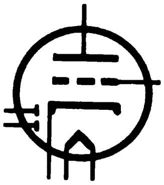
ELECTRON-RAY INDICATOR TUBE



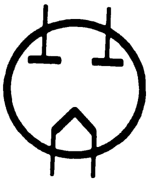
KLYSTRON



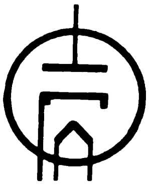
DIODE, DIRECTLY HEATED CATHODE



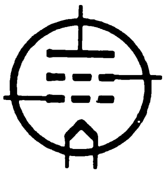
DUPLEX-DIODE TRIODE



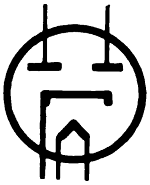
DUO-DIODE, DIRECTLY HEATED CATHODE



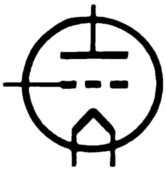
DIODE, INDIRECTLY HEATED CATHODE



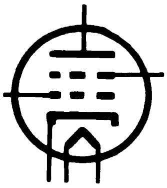
TETRODE, DIRECTLY HEATED CATHODE



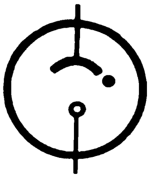
DUO-DIODE, INDIRECTLY HEATED CATHODE



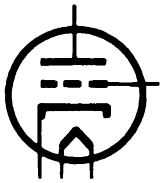
TRIODE, DIRECTLY HEATED CATHODE



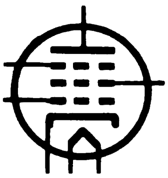
TETRODE, INDIRECTLY HEATED CATHODE



VOLTAGE REG, COLD CATHODE—GASEOUS

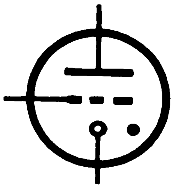


TRIODE, INDIRECTLY HEATED CATHODE



PENTODE, INDIRECTLY HEATED CATHODE

BEAM-POWER AMPLIFIER



COLD CATHODE TRIODE—GASEOUS

Index

A

Absolute value 508
AC, defined 82
Acorn tubes 184, 339
AC plate resistance 163
AC values
  average 85
  effective 86
  instantaneous 85
  maximum or peak 85
AFC 450
AFC sweep generator 450
Algebra 507-516
Alternation 84
Alternator 82
Altitude delay 460, 465
Ammeter 20, 73, 79
Amplification factor 161, 163
Amplifiers 215-232
Amplitude 85
Amplitude modulation 252, 261, 263
AN/APN-9 202, 250
AN/APS-23
  AC servo loop 495
  antenna indicator servo system 493
  delay circuit 468
  indicator 458
  range marker circuit 463
  transmission channel 438
AN/APS-42
  delay circuit 470
  transmission channel 434
Antennas 269, 411-432
  arrays 425
  current-fed 421
  directivity 412
  efficiency 412
  half-wave 418
  impedance 418, 419
  reciprocity 411
  voltage-fed 421
Anti-jam switch 448
Anti-TR circuit 403
Apparent power 121, 129, 131, 134
Aquadag 192
Artificial magnets 43
Artificial transmission lines 376-380
A-scan 327
AS-361 430
Attenuation 358, 392
ATR devices 407
Audio amplifiers 222
Audio amplifier system 280
Audio power amplifier 260, 280
Audio transformers 112
Automatic frequency control 450
Autotransformer 113
Average power 331
Azimuth stabilization 493

B

Back EMF 64, 68, 88
Bandwidth 142, 147, 223, 226, 228
Bandwidth control 446
Battery connections 24
Battery, storage 22
Beam power tubes 180
Bends in waveguides 397
Bias 161, 165, 170
Biased limiters 318
Biased restorers 321
Bleeder resistor 210
Blocked oscillator 305
Blocking oscillator 301
Bridge rectifiers 206
B-scan 327
Buffer amplifiers 258
Bunching, electron 348

C

Capacitance, formulas 102
Capacitance, interelectrode 173
Capacitive reactance 102
Capacitor-input filter 209
Capacitors 99-106
Carbon microphone 259
Cascade, amplifiers in 222
Cathode bias 170
Cathode follower 231
Cathode ray tubes 189-196
Cathodes 153
Cavity resonators 400-410
Characteristic 519
Characteristic curves 162, 175, 177, 184
Characteristic impedance 355, 357, 391
Charged bodies 5, 7
Chemical effect of current 69
Choke-input filters 210
Choke joint 398
Choke, RF 97
Chokes, filter 95
Circuit breakers 50
Circuit, closed 17
Circuit, open 17
Clampers
  biased 321
  negative 320
  positive 321
Class A amplifiers 216
Class B amplifiers 217
Class C amplifiers 219
Clippers 316
Closed circuit 17
Close-wound coils 89
Coaxial line 371
Coefficient of coupling 109
Coil, laminated iron-core 96
Color code, resistor 12
Colpitts oscillator 243
Common logarithms 519
Communication system 197
Compensation, high frequency 225
Compensation, low frequency 225
Complex numbers 518
Compound motor 65
Compound-wound generators 59
Concentric line 371
Conductance 9
Conductivity 10
Contact coupling 397
Contrast circuit 446
Contrast control 446
Control ratio 188
Controls, linear 41
Controls, tapered 41
Control transformer 486
Copper losses 67
Copper-oxide rectifier 77
Cosecant-squared reflector 428
Counter EMF 64, 88
Counting circuit 307
Coupling
  direct 220
  impedance 221
  RC 220
  transformer 221
CRT, electromagnetic 193
CRT, electrostatic 189
Crystal axes 245
Crystal-controlled oscillators 248
Crystal cuts 245
Crystal holders 247
Crystal, temperature coefficient 246
Crystal diode limiter 318
Crystal diodes 158
C-scan 329
Current-fed antenna 421
Cutler feed 429
Cutoff frequency 389, 390
Cycloids 342

D

Damping 72, 235
D'Arsonval meter 70
DC amplifiers 222
DC generator 57
DC motor 63
DCR 320
DC resistance of vacuum tubes 164
Deflection coils
  factor 193
  plates 192
  sensitivity 193
Delay circuits 465
Detection 271, 277
Development of radar 1-2
Diamagnetic substances 46
Dielectric absorption 105
Dielectric 99, 101, 105
Dielectric hysteresis 105
Dielectric losses 104
Differential generator 484
Differentiation 295, 296
Diode detector 278
Dipole, folded 427
Direct coupling 220
Direct-current restorer 320
Directivity 412, 426
Directly-heated cathodes 153
Discriminator 450
Distributed constants 355, 356
Double diode limiters 319
Double-tuned amplifiers 227
DR computer 118, 120, 122, 123, 126, 128, 130, 133, 525, 532
Driver 259
Dry cell 21
Duplexing system
  AN/APS-23 407-410
  Parallel 403
  Series-parallel 405
Dynatron 175
Dynamic curve 166

E

Eccles-Jordan multivibrator 312
Eddy current losses 68, 113
Edison effect 149
Electrical zero 477, 487
Electrolytic capacitors 106
Electromagnetic CRT 193, 195
Electromagnetic energy in space 415
Electromagnetic fields 357, 386
Electromagnetic wave motion 413
Electromagnets 48
Electron bunching 348
Electron emission 151
Electron flow, producing 15
Electron gun 189
Electronic voltage regulators 212
Electroscope 7
Electrostatic CRT 189, 191
Electrostatic fields in diode 155
Electrostatic fields in triode 159
EMF, counter 64
EMF, induced 51
Energy 28
Extinction potential 186

F

Feedback, degenerative 234
Feedback, regenerative 234
Ferromagnetic substances 46
Field emission 151
Field intensity 46
Fields of force 6
Fighter radar 334
Filter chokes 95
Filter, frequency rejection 148
Filters 207-211
Firing point 185, 188
Fixed bias 170
Flange-to-flange coupling 397
Flat line 366
Fleming's valve 150
Flexible waveguides 399
Fluorescent screen 192
Flux density 46
Flux linkages 55
Focus control 194
Focusing coil 193
Folded dipole 427
Force 28
Force, fields of 6
Form 781 write-ups 506
Fractional equations 514
Fractional exponents 516
Fractions 512

- Frequency, audio 82
  - power 82
  - radio 82
- Frequency conversion 275
- Frequency divider 306
- Frequency modulation 252
- Frequency
  - of oscillation 135, 136, 243, 244, 245
- Frequency rejection filter 148
- Front-to-back ratio 426
- Full-wave rectifier 204
- Fuses 26

G

- Gain control 228
- Gas tubes 184
- GEE 335
- Generators
  - AC 56
  - compound-wound 59
  - DC 57
  - load characteristics 60
  - self-excited 58
  - series-wound 58
  - shunt-wound 59
- Getter 155
- Grid-leak bias 171, 238
- Grid modulation 264
- Ground 17
- Group velocity 391

H

- Half-wave rectifier 204
- Hard-tube sweep generator 315
- Hartley, modified 242
- Hartley oscillator 241, 251
- Hartley, shunt-fed 243
- Height-finding radar 334
- Heterodyning 261, 273
- High frequency compensation 225
- Hydrometer 23
- Hyperbolic sweep 460
- Hysteresis 68, 113

I

- Identification 334
- IF amplifiers 227
- IFF 334
- Image frequency 274
- Imaginary numbers 517
- Impedance 87, 117, 121, 126, 139
  - antenna 418, 419
  - coupling 221
  - line 362, 368
  - transformers 374, 396
- Indicator AC servo loop 495
- Indirectly-heated cathodes 153
- Induced EMF 51
- Inductance 88
- Induction fields 415
- Induction, self 54
- Inductive reactance 91
- Inductor, phase relationship 91
- Inductors 87-98
- Inductor, variable 98
- Instantaneous values 85
- Integration 295, 296
- Intensity modulation 460
- Interlectrode capacitance 173, 337
- Interpolation 520
- Inverse peak voltage 157
- Ionization 5
- Ions 5
- Iris 394
- Iron losses 68

K

- Kirchhoff's laws 30

L

- Laminated iron core 96
- Layer-wound coils 89
- Leakage in dielectric 104
- Leakage flux 109
- Lenz' law 53
- LC filters 209
- Lighthouse tube 339
- Limiters
  - biased 318
  - crystal diode 318
  - double diode 319
  - series diode 317
  - shunt diode 317
  - triode 317
- Linear controls 41
- Load 38, 40
- Load lines 169

Local oscillator 276, 440
Locked-in sweep 452
Lodestone 43
Logarithms 519
Long time-constant 296
Loop 394, 402
Loran 202, 214
Losses
  copper 67
  eddy current 68
  hysteresis 68
  in inductors 94
Low frequency compensation 225
L-type filter 209

M

Magnetic fields 45-48
Magnetic shielding 47
Magnetism, theory of 47
Magnetization curve 96
Magnetron 340-348
Magnets 43
Mantissa 519
Matter 3
Master oscillator power amplifier 256
Mechanical axes 245
Metallic insulator 372
Meter, D'Arsonval 70
Meter damping 72
Meter sensitivity 72, 76
Mica dielectric 105
Microphone, carbon 259
Microwave oscillators 337-352
Microwave wavemeter 402
Mixers 275
Modes
  in magnetron 346
  in waveguides 389
  in reflex klystron 350
Modulation 198, 252, 254, 261, 263
Modulation factor 261
Modulator 259
MOPA 256, 257, 258
Motor
  compound 65
  series 64
  shunt 65
Multielectrode tubes 180
Multimeter 77
Multiunit tubes 181
Multivibrators
  Eccles-Jordan 312
  plate-coupled 309
  start-stop 311
Mutual characteristic curves 162
Mutual conductance 163
Mutual inductance 90, 108

N

Natural logarithms 519, 521
Negative resistance 175
Neutralization 256
Neutron 4
Nonresonant lines 366, 423
Nucleus 5

O

Ohmmeter 20, 76
Ohm's law 19
Oil dielectric 105
Open circuit 17
Optical axis 245
Oscillations in tank 234
Oscillators 233-250

P

Paper dielectric 105
Parabolic reflectors 427
Parallel resonance 144
Paramagnetic substances 46
Parasitic arrays 425
Peaked wave 284
Peak power 331
Peak value 85
Pedestal, unblanking 448
Pentagrid converter 181, 276
Period 84
Permeability 46
Persistence of screen 192
Phase angle 118, 121, 123, 125, 130, 133
Phase detector 500
Phase relationship
  in capacitor 102
  in inductor 91
Phase velocity 391
Photoelectric cell 15, 16
Photoelectric effect 16, 70
Photoelectric emission 152

Piezoelectric effect 70, 245
Pi-type filter 209
Planetary electrons 4
Plan position indicator 329
Plate-coupled multivibrator 309
Plate modulation 254, 263
Polarity 30
Polarization 21, 420
Power 28, 35
Power amplifier 215, 257, 260, 280
Power losses
  copper 67
  eddy current 68
  hysteresis 68
  in capacitors 104
  in inductors 94
Power rating 39
Powers of ten 527-530
PPI scan 329
Preamplifiers 446
PRF 330
Principles of radar 323, 326
Probe coupling 394, 402
Producing electron flow 15
Pulse repetition frequency 330
Pulse shape 329
Pulse system 325, 326
Pulse width 330
Push-pull amplifiers 218

Q

Q of tuned circuit 142, 147, 228, 236
Quadratic equations 518

R

Radar, development of 1-2, 332
Radar principles 323-336
Radiation fields 416, 424
Radicals 516
Radio compass 201
Range-marker circuit 463
Range resolution 330
RC circuits 121-124, 130-132, 285
RC coupling 220, 223
RC time constants 286
RCL circuits 124-126, 132-134
Reactors, saturable 96
Recovery time 406
Rectifier, copper-oxide 77
Rectifiers 157, 203-214
Reflected waves 360
Reflectors, cosecant-squared 428
Reflectors, parabolic 427
Reflex klystron 348-352
Regulation, voltage 59
Regulator tubes 186
Relays
  overload 50
  starting 49
Remote-cutoff tubes 179
Resistance 9
Resistance, negative 175
Resistor color code 12
Resistor, power rating 39
Resistors 10-11
Resonance 135-148
Resonance curve 138
Resonant cavity 401
Resonant frequency 135, 136, 245
Resonant lines 366, 422
RF amplifiers 226
RF choke 97
Rheostats 11, 40
RL circuits 116-121, 128-130, 297
Rotating joints 398

S

Saturable reactors 96
Sawtooth wave 284
Scan types 327
SCR 718 altimeter 201
Screen, fluorescent 192
Secondary emission 151, 175
Selectivity 142, 147, 268, 270
Self-excited oscillator 238
Self-inductance 90
Sensitivity, deflection 193
Sensitivity, meter 72, 76
Sensitivity, receivers 267
Series diode limiter 317
Series motor 64
Series-peaking coil 226
Series resonance 136
Series-type ohmmeter 76
Series-wound generator 58
Servo loop 495
Servomechanisms 471, 489-493
Servo system, AN/APS-23 493
Shepherd-Pierce tube 352

Shielded-pair line 372
Short time-constant 297
Shunt diode limiter 317
Shunt, meter 73
Shunt motor 65
Shunt peaking coil 226
Shunt-type ohmmeter 76
Shunt-wound generator 59
Sidebands 261, 277
Signal-to-noise ratio 446
Simultaneous linear equations 515
Single-tuned amplifiers 227
Single-wire line 370
Slip rings 82
Slot 394
Space charge 154
Space current 161
Space-wound coils 89
Square wave 284, 290-294, 299
Squares, cubes, and roots 551, 552
Stagger tuning 228
Standing-wave ratio 365
Standing waves 360, 361, 363, 364
Start-stop multivibrator 311
Step-charging diode 306
Step-down transformer 110
Step-up transformer 110
Striking potential 185
Stub 373, 396
Stub matching 373
Structure, atom 4
Structure, matter 3
Supercontrol tubes 180
Superheterodyne receiver 272, 282
Surface barrier 152
Surge impedance 355
Sweep generators 314, 350
  hard tube 315
  thyration 314
Switches 26
Synchro 471-489
Synchro differential 482
Synchro-differential motor 486
Synchronizing a blocking oscillator 305
Synchro schematics 475

T

Tantalum 152
Tapered controls 41
Tapered-line transformer 396
TE 389
Technical Orders 505, 506
Teeing 373
Temperature coefficient of crystal 246
Temperature saturation 155
Ten, powers of 527-530
Testing capacitor 106
Testing inductors 98
Testing transformers 114
Thermocouple 78
Thermionic emission 152
Thoriated-tungsten cathodes 153
Thyratrons 187
Thyratron sweep generator 314
Tickler-coil oscillator 237
Time-base generator 314
Time constants 286, 296
TM 389
Torque 62
TPTG Oscillator 244
Transformer coupling 221
Transformers 107-114
  impedance ration 111
  power 111
  step-down 110
  step-up 110
  turns ratio 109
Transistor 182
Transit time 183, 338
Transmission lines 353-380
Transverse electric (TE) 389
Transverse magnetic (TM) 389
Traveling waves 360, 362, 363
TR circuit 403
Trigonometric functions 521
Triode limiters 319
TR tubes 405
True power 121, 129, 131, 134
Tuned circuit Q 142
Tuned-grid oscillator 240
Tuned line 368
Tuned-plate oscillator 240
Tuned-plate, tuned-grid oscillator 244
Tungsten 152
Tuning screw 395
T-type filter 209
Twisted-pair line 372
Two-wire parallel line 370

U

UHF tubes 339
Universal shunt 73
Universal Time Constant Chart 288
Untuned line 366

V

Variable-mu tubes 180
Variac 114
Vector diagram analysis 139
Vectors 533
Video amplifiers 224
Video circuit 448
Voltage
  amplifiers 215, 260, 280
  dividers 38
  doublers 206
  gain 142
  phase relationships 110
  ratings 105
  regulation 59
  regulators 211
Voltage-fed antennas 421
Voltmeter 20, 74, 79
Voltmeter, sensitivity 76
Volume control 215
VR tubes 211

W

Waveguides 381-400
  attenuation 392
  boundary conditions 384
  characteristic impedance 391
  coupling 393
  cutoff frequency 389, 390
  energy flow 386
  flexible 399
  group velocity 391
  impedance matching 395
  modes 389, 392
  phase velocity 391
  power handling capabilities 392
  wavelength 391
Wavelength 84, 355, 391
Wavemeter, microwave 402
Wave motion on transmission lines 359
Work 28
Work function 152
Window coupling 394, 402
Window reactances 396

XYZ

X-axis 245
Y-axis 245
Z-axis 245

☆ U.S. GOVERNMENT PRINTING OFFICE: 1963—651239/24

Date Due

MAY 20 1983
DEC 22 1984
NOV 26 '88
JAN 27 '92

UG633
.A3763
no.101-8

c.1

United States Department of the Air
Force

Fundamentals of electronics

Google

A000027897041